



## Accounting for imperfect forward modeling in geophysical inverse problems — Exemplified for crosshole tomography

Hansen, Thomas Mejer; Cordua, Knud Skou; Holm Jacobsen, Bo; Mosegaard, Klaus

*Published in:*  
Geophysics

*Link to article, DOI:*  
[10.1190/GEO2013-0215.1](https://doi.org/10.1190/GEO2013-0215.1)

*Publication date:*  
2014

*Document Version*  
Publisher's PDF, also known as Version of record

[Link back to DTU Orbit](#)

*Citation (APA):*  
Hansen, T. M., Cordua, K. S., Holm Jacobsen, B., & Mosegaard, K. (2014). Accounting for imperfect forward modeling in geophysical inverse problems — Exemplified for crosshole tomography. *Geophysics*, 79(3), H1-H21. <https://doi.org/10.1190/GEO2013-0215.1>

---

### General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

# Accounting for imperfect forward modeling in geophysical inverse problems — Exemplified for crosshole tomography

Thomas Mejer Hansen<sup>1</sup>, Knud Skou Cordua<sup>1</sup>, Bo Holm Jacobsen<sup>2</sup>, and Klaus Mosegaard<sup>1</sup>

## ABSTRACT

Inversion of geophysical data relies on knowledge about how to solve the forward problem, that is, computing data from a given set of model parameters. In many applications of inverse problems, the solution to the forward problem is assumed to be known perfectly, without any error. In reality, solving the forward model (forward-modeling process) will almost always be prone to errors, which we referred to as *modeling errors*. For a specific forward problem, computation of crosshole tomographic first-arrival traveltimes, we evaluated how the modeling error, given several different approximate forward models, can be more than an order of magnitude larger than the measurement uncertainty. We also found that the modeling error is strongly linked to the spatial variability of the assumed velocity field, i.e., the a priori velocity model. We discovered some general tools by which the modeling error can be quantified and cast into a

consistent formulation as an additive Gaussian observation error. We tested a method for generating a sample of the modeling error due to using a simple and approximate forward model, as opposed to a more complex and correct forward model. Then, a probabilistic model of the modeling error was inferred in the form of a correlated Gaussian probability distribution. The key to the method was the ability to generate many realizations from a statistical description of the source of the modeling error, which in this case is the a priori model. The methodology was tested for two synthetic ground-penetrating radar crosshole tomographic inverse problems. Ignoring the modeling error can lead to severe artifacts, which erroneously appear to be well resolved in the solution of the inverse problem. Accounting for the modeling error leads to a solution of the inverse problem consistent with the actual model. Further, using an approximate forward modeling may lead to a dramatic decrease in the computational demands for solving inverse problems.

## INTRODUCTION

Computation or prediction of geophysical data  $\mathbf{d}$  given a set of model parameters  $\mathbf{m}$  which represents a model of the subsurface, is referred to as *solving the forward problem*. In mathematical terms, the forward problem is defined through a, possibly nonlinear, operator  $\mathbf{g}$  that relates model parameters  $\mathbf{m}$  to data  $\mathbf{d}$  such that solving the forward problem can be given by

$$\mathbf{d} = \mathbf{g}(\mathbf{m}). \quad (1)$$

The associated inverse problem deals with the problem of inferring information about the subsurface, through the model parameters  $\mathbf{m}$  given a set of observed data  $\mathbf{d}_{\text{obs}}$  and the forward model  $\mathbf{g}$ . The

accuracy of the forward model directly influences the solution to the inverse problem.

In many important cases, only a more or less accurate approximation to  $\mathbf{g}(\mathbf{m})$  is available. This introduces an error in the forward model that will lead to erroneous data calculations. We refer to this type of error as a modeling error, which is also referred to as the *theoretical error* (Tarantola and Valette, 1982a) or the *modelization error* (Tarantola, 2005). The modeling error refers to any error related to solving the forward problem and can be categorized into at least the following four categories: (1) numerical precision, (2) choice of physical model, (3) approximate solution of the physical problem, and (4) incomplete model description.

- 1) Numerical precision — Computers represent numbers with a certain precision. Therefore, any computation on a computer

Manuscript received by the Editor 11 June 2013; revised manuscript received 11 December 2013; published online 13 March 2014.

<sup>1</sup>Technical University of Denmark, DTU Space, National Space Institute, Department of Mathematical and Computational Geoscience, Lyngby, Denmark. E-mail: tmeha@dtu.dk; kcor@dtu.dk; kmos@imm.dtu.dk.

<sup>2</sup>Aarhus University, Department of Geoscience, Aarhus C, Denmark. E-mail: bo@geo.au.dk.

© 2014 Society of Exploration Geophysicists. All rights reserved.

- will be accurate only to a certain precision. In addition, any solution to a partial differential equation, using e.g., finite-difference or finite-element solutions, relies on choices of discretization that affect the accuracy of the numerical solution.
- 2) Choice of physical model — A complete description of the physical problem is often not computationally viable. Seismic waves, for example, propagate in a 3D viscoelastic anisotropic medium. Although methods exist to simulate waveform propagation in such a medium (Saenger and Bohlen, 2004), it may be computationally too expensive, or complex, to apply for some problems. Instead, solving the isotropic elastic or acoustic wave equation may be faster, but it will also introduce a modeling error. As another example, which we investigate in more detail later, different physical theories (linear and nonlinear) have been used to compute first-arrival traveltimes, used in, for example, crosshole tomography (see, e.g., Vidale, 1988; Cervený and Soares, 1992; Jensen et al., 2000; Spetzler and Snieder, 2004). In this case, a modeling error typically will be introduced due to the choice of physical model describing the forward model.
  - 3) Approximate solution of the physical problem — A common type of modeling error stems from the fact that geophysical data are responses of a 3D earth. A full 3D solution of the forward problem may be computationally too expensive to apply for some problems. The inverse problem may be solved based on faster 1D, 2D, or 2.5D forward models. For example, massive amounts of 3D airborne electromagnetic data have been inverted using least-squares-type inversion based on a computationally very efficient 1D forward model (see, e.g., Viezzoli et al., 2008). Likewise, Minsley (2011) make use of an efficient 1D forward model in a probabilistic inversion of airborne electromagnetic data. Seismic data can in principle be inverted based on accurate 3D full-waveform (FW) modeling (see, e.g., Tarantola, 1988). However, due to computational demands, inversion of reflection seismic data often relies on faster 1D forward modeling codes (see, e.g., Buland and Omre, 2003).
  - 4) Incomplete model description — Yet another type of modeling error is linked to the parameterization of the inverse problem and description of the physical system in which data are recorded. Christiansen et al. (2011) analyze errors due to an imperfect system description related to 1D transient electromagnetic forward modeling and find that such errors can lead to an error in the estimate of the subsurface layer resistivities that is an order of magnitude of the true layer resistivities. Cordua et al. (2008, 2009) discuss and demonstrate how borehole cavities can lead to significant modeling errors in crosshole georadar tomography if not accounted for.

Several studies have acknowledged and exemplified that the choice of forward-modeling code may not only affect the forward modeling, but also the result of the associated inverse problem. For example, Cuffin (1981) analyzes the effects related to using an approximate forward model related to least-squares-type inversion of electrocardiology data. Lelièvre et al. (2011) analyze the modeling error related to first-arrival computation between two different forward models and show examples of deterministic inversion based on both forward models. Lekić and Romanowicz (2011) describe how different sources of Gaussian modeling errors related to

crustal size waveform tomography result in errors larger than typical measurement uncertainties. Moreover, they discuss that the existence of a bias in the modeling error may cause the location of the model with apparent maximum likelihood to be displaced from the model with actual maximum likelihood. Although the existence of the modeling error related to geophysical inverse problems is acknowledged, no formal way of characterizing, and accounting for, the modeling error in a probabilistic manner is currently known.

This study provides some general tools by which the modeling error can be quantified and cast into a consistent formulation as an additive Gaussian observation error. The theory and methodologies that follow apply to any inverse problem. In the remainder of this paper, we will though, as an example, consider the modeling error related to cross-borehole ground-penetrating radar (GPR) tomography.

First, we demonstrate that the use of an approximate method for solving the forward problem, of computing the traveltime delay between a source and a receiver, can lead to modeling errors that are an order of magnitude larger than typical measurement uncertainty. We also demonstrate that the magnitude of the modeling error is closely linked to the type of subsurface heterogeneity.

Then, we propose a method for sampling (that is, providing several realizations of) the modeling error related to the difference between two different forward models. A probabilistic model of the modeling error is inferred in the form of a correlated Gaussian probability distribution with mean  $\mathbf{d}_T$  and covariance  $\mathbf{C}_T$ ,  $\mathcal{N}(\mathbf{d}_T, \mathbf{C}_T)$ . This allows accounting for the modeling error as an additive Gaussian observational error.

Finally, we show examples based on crosshole tomography that demonstrate the effect of disregarding and accounting for the modeling error, respectively.

## MODELING ERRORS IN CROSS-BOREHOLE GPR TOMOGRAPHY

Cross-borehole tomography is a method widely used in geophysical prospecting for characterizing small-scale variations of near-surface environments. Often, the tomographic images are based on first-arrival traveltimes of seismic or electromagnetic signals that are propagated between the boreholes. In this study, we consider an example from GPR cross-borehole tomography. This method has become popular during the past few decades and has various applications, such as mapping of tunnels and voids (Moran and Greenfield, 1993), mapping of bedrock fractures and fracture zones (Olsson et al., 1992; Lane et al., 1998), estimation of hydrological parameters and delineation of flow paths in the unsaturated zone (Hubbard et al., 1997; Looms et al., 2008a, 2008b), and delineation of geologic structures and lithologies (Fullagar et al., 2000; Bellefleur and Chouteau, 2001; Tronicke et al., 2004).

Several sources of modeling error associated with crosshole georadar tomography have been discussed. Peterson (2001) and Cordua et al. (2008) list errors introduced in GPR cross-borehole tomography caused by incorrect station geometry, zero time calibration, geometric spreading, transmitter radiation pattern, transmitter amplitude, and high angle raypaths. Cordua et al. (2009) demonstrate that some of these errors (related to, e.g., cavities in the borehole walls) result in correlated data errors. They also demonstrate how a model of Gaussian correlated data errors that accounts for these errors can be empirically set up and used in a probabilistic inversion approach. Another type of modeling error, which was

not described by Peterson (2001) and Cordua et al. (2008, 2009), is associated with the way the forward problem of computing the traveltime between a source and a receiver is solved. We will consider this type of modeling error in more detail.

### Forward models for first-arrival traveltime computation

The forward problem related to traveltime tomography consists of computing the time between emitting a seismic or electromagnetic waveform at a source location and the first-arrival time of the propagating waveform recorded at a receiver. By first arrival, we specifically mean the first-break arrival time. Many methods exist to solve the forward problem. Here, we briefly discuss some of the most widely used methods to solve the forward problem related to traveltime computation. Subsequently, we will refer to a specific choice of forward model using a subscript to  $\mathbf{g}$  as defined in the following.

The forward models defined below will be discussed and compared by considering three different reference velocity models of size  $6 \times 3$  m, with a source located at  $S = (0.5 \text{ m}, 1.5 \text{ m})$  and a receiver located at  $R = (5.5 \text{ m}, 1.5 \text{ m})$  (see Figure 1, top).

Model a is a constant velocity model with a velocity of 0.14 m/ns (Figure 1a, top). Model b is a realization of a Gaussian random field with mean velocity of 0.14 m/ns, variance of  $1.44e-4$  (m/ns)<sup>2</sup>, and an exponential covariance model with a direction of maximum continuity of 10° below horizontal, a maximum correlation length of 20 m, and a minimum correlation length of 2 m (see Figure 1b, top). Model c is a realization of a random function inferred from a binary-channel-based training image, generated using single normal equation simulation (SNESIM), Strebel (2002). The velocity outside the channels is 0.1273 m/ns, and it is 0.1673 m/ns inside the channel. The mean velocity is 0.14 m/ns (see Figure 1c, top).

The first-arrival sensitivity kernels corresponding to each of the considered forward models for each of the three considered reference models are computed using SIPPI (Hansen et al., 2013b) and visualized below each velocity model in Figure 1, where black indicates positive sensitivity (an increase in velocity will cause a decrease in the traveltime) and red indicates negative sensitivity. The sensitivity kernel is the first-order Fréchet derivative of the forward model with respect to a given velocity model.

#### High-frequency (ray) approximation ( $\mathbf{g}_{\text{SR}}, \mathbf{g}_{\text{BR}}$ )

The simplest forward model for computing first-arrival traveltimes relies on a high-frequency approximation of the wave equation, often referred to as the *ray* approximation. The traveltime is found as the delay caused by the signal traveling along the fastest raypath connecting a source and a receiver. This traveltime can be

efficiently computed using the eikonal solution to the wave equation (e.g., Vidale, 1988). We will refer to this as the *bending ray* forward model  $\mathbf{g}_{\text{BR}}$ , which is nonlinear because the raypath depends on the velocity model. A further simplification is to assume that the raypath follows a straight line between the source and receiver. This leads to the linear straight-ray forward model  $\mathbf{g}_{\text{SR}}$ , which has probably been the most widely used forward model in traveltime tomography (e.g., Cordua et al., 2008). The model  $\mathbf{g}_{\text{SR}}$  is identical to  $\mathbf{g}_{\text{BR}}$  in case the velocity field is homogeneous. The second and third rows in Figure 1 show the sensitivity kernels related to the  $\mathbf{g}_{\text{SR}}$  and  $\mathbf{g}_{\text{BR}}$  forward models for the three reference models.

#### Finite-frequency (fat-ray) approximation ( $\mathbf{g}_{\text{SFR}}, \mathbf{g}_{\text{BFR}}$ )

The ray approximation has become popular because it is simple and computationally cheap. However, the frequency of a propagating wave is always band limited, which results in scattering effects that are not accounted for by the ray approximation. Therefore, so-called fat rays have been considered. For fat rays, the traveltime is sensitive not only to the traveltime delay along the fastest raypath,

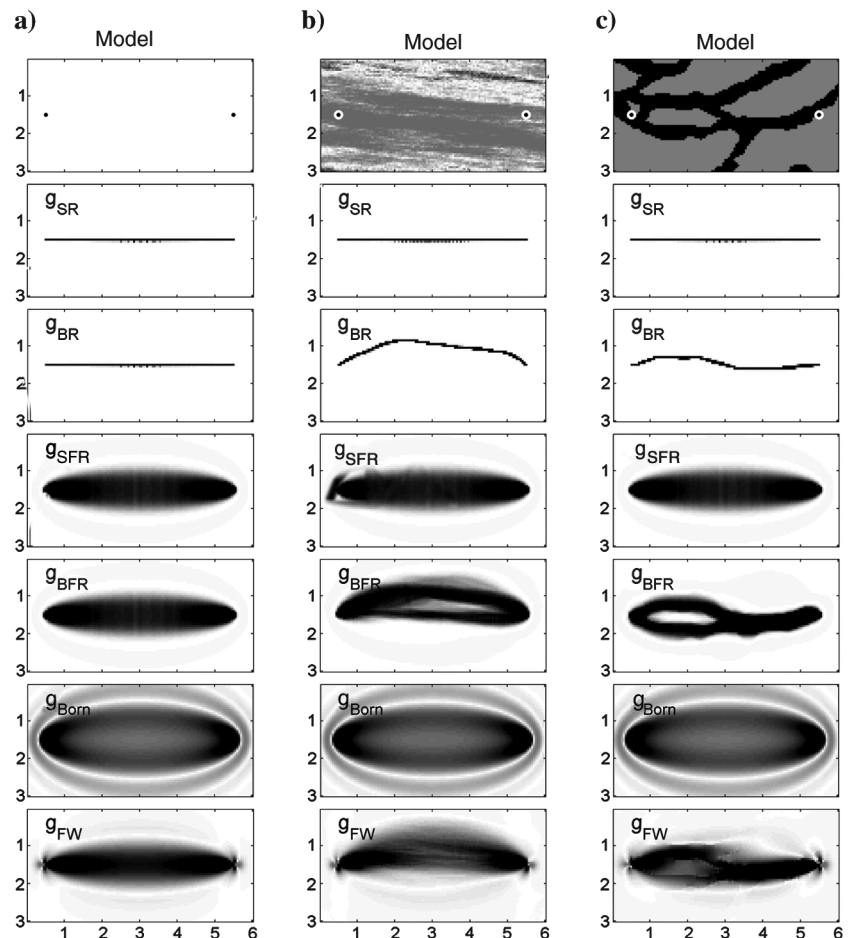


Figure 1. Top row: reference velocity models (a) homogeneous, (b) realization of a Gaussian model, and (c) realization based on multiple point statistics inferred from a training image. White reflects a high velocity, and black reflects a low velocity. Each column shows the sensitivity kernels associated with the forward models  $\mathbf{g}_{\text{SR}}, \mathbf{g}_{\text{BR}}, \mathbf{g}_{\text{SFR}}, \mathbf{g}_{\text{BFR}}, \mathbf{g}_{\text{Born}}$ , and  $\mathbf{g}_{\text{FW}}$ . For the sensitivity kernels, black indicates positive sensitivity and white indicates negative sensitivity.



but to an area around the raypath (see, e.g., Marquering et al., 1999; Dahlen et al., 2000; Jensen et al., 2000; Spetzler and Snieder, 2004; Buursink et al., 2008). For a single-frequency source wavelet, the sensitivity kernel consists of alternating regions of positive and negative sensitivity, known as *Fresnel zones* (Woodward, 1992). Most of the sensitivity for a band-limited source lies within the first Fresnel zone (Jensen et al., 2000) because higher order Fresnel zone sensitivity tend to cancel out. Several empirical models have been suggested that describe a sensitivity kernel based on the first Fresnel zone (Cerveny and Soares, 1992; Jensen et al., 2000; Husen and Kissling, 2001). Here, we specifically make use of the description proposed by Jensen et al. (2000) to compute a sensitivity kernel reflecting the first Fresnel zone. We will refer to this nonlinear kernel, where the sensitivity is dependent on the velocity model, as the *bending Fresnel* forward model  $\mathbf{g}_{\text{BFR}}$ . We will also consider a related linear kernel, obtained by assuming a constant velocity field, and we refer to this as the *straight Fresnel* forward model,  $\mathbf{g}_{\text{SFR}}$ . The fourth and fifth rows in Figure 1 show the sensitivity kernels related to the  $\mathbf{g}_{\text{SFR}}$  and  $\mathbf{g}_{\text{BFR}}$  forward models for the three reference models.

#### The Born approximation ( $g_{\text{Born}}$ )

Using the Born approximation (considering first-order scattering), an exact analytical expression for the sensitivity kernel for a point source can be derived for seismic (Marquering et al., 1999; Dahlen et al., 2000; Jensen et al., 2000; Spetzler and Snieder, 2004; Liu et al., 2009) and electromagnetic wave propagation (Buursink et al., 2008; Liu et al., 2009). Here, we will make explicit use of the formulation of the sensitivity kernels given by Buursink et al. (2008), and we refer to it as the *Born* forward model  $g_{\text{Born}}$ .

The sixth row in Figure 1 shows the sensitivity kernels related to the  $g_{\text{Born}}$  forward model for the three reference models. Note that, as opposed to the previously considered sensitivity kernels, regions of positive and negative sensitivity are visible. The Born approach is only strictly valid for a homogeneous velocity model, and while sensitivity kernels in principle can be computed for velocity models with small velocity contrasts, the Born approach will fail for larger velocity contrasts. We therefore only consider the sensitivity kernel associated with a homogeneous velocity model. This type of forward model is then linear for a constant velocity, independent of the actual velocity variations. Note also that sensitivity kernels associated with the Born approximation typically do not reflect the sensitivity of the first-break arrival (as we make use of to determine the traveltime in this paper) but the time delay associated with maximum crosscorrelation between the observed and simulated wavefield. We still consider the sensitivity kernels based on the Born approximation here because they have previously been used to invert traveltime data recorded as first-break arrival (see e.g., Buursink et al., 2008; Liu et al., 2009).

#### Waveform modeling and first-arrival picking ( $\mathbf{g}_{\text{FW}}$ )

Perhaps the most precise, and time consuming, approach to solve the forward problem is to use FW modeling, followed by picking the first-arrival time. We refer to such a forward model as the *FW forward model*  $\mathbf{g}_{\text{FW}}$ . Specifically, we use 2D finite-difference waveform modeling (Ernst et al., 2007) with a Ricker source wavelet with a peak frequency of 100 MHz. First-break arrival times are automatically picked using the method proposed by Molyneux

and Schmitt (1999). The first-order Fréchet derivative, i.e., the sensitivity kernel, related to the  $\mathbf{g}_{\text{FW}}$  forward model is estimated using the perturbation approach (McGillivray and Oldenburg, 1990). Here, each model parameter is perturbed slightly and the resulting residual traveltime is compared to the traveltime related to the reference model. From this, the local gradient, i.e., the first-order Fréchet derivative, can be computed. The resulting sensitivity kernel obtained using the perturbation approach for the three considered reference models is shown in the bottom row in Figure 1.

Even if the  $\mathbf{g}_{\text{FW}}$ -type forward model is the most precise way of computing traveltimes we consider here, it is in itself associated with modeling errors as compared to actual observed traveltimes. The finite-difference-based waveform modeling code is 2D isotropic and is second-order accurate in space and time. One could choose a 3D anisotropic code for increased accuracy and higher order finite-difference schemes for improved modeling accuracy.

Note that the sensitivity kernels associated with the  $\mathbf{g}_{\text{SR}}$ ,  $\mathbf{g}_{\text{BR}}$ ,  $\mathbf{g}_{\text{SFR}}$ ,  $\mathbf{g}_{\text{BFR}}$ , and  $g_{\text{Born}}$  forward models completely define the forward model with respect to the given reference models because the full Fréchet derivatives are completely given by the first-order Fréchet derivative. This is not the case for the  $\mathbf{g}_{\text{FW}}$  forward model, which is sensitive to higher order scattering. The sensitivity kernels shown in the bottom row in Figure 1 are only the first-order Fréchet derivatives, and hence only sensitive to first-order scattering.

From Figure 1, it is evident that the sensitivity kernels differ, based on the underlying assumptions. The fat ray ( $\mathbf{g}_{\text{SFR}}$ ,  $\mathbf{g}_{\text{BFR}}$ ) and Born-( $g_{\text{Born}}$ ) forward models only consider first-order scattering, although in reality, higher order scattering may occur. The high-frequency forward models ( $\mathbf{g}_{\text{SR}}$ ,  $\mathbf{g}_{\text{BR}}$ ) do not consider scattering effects at all. Using any of these considered forward models will result in a possible different traveltime than that obtained through the finite-difference calculation.

Other forward models for traveltime computation can also be examined (e.g., Tromp et al., 2005; Lelièvre et al., 2011). Our goal here is not to discuss the validity of each type of forward model, but to describe that using any type of forward model will lead to different modeling errors affecting the computed traveltimes.

### Modeling errors in first-arrival traveltime computation

We will now analyze the modeling error caused by using a specific choice of forward model. The effect of different kinds of variability in the model space is investigated based on 26 reference models that are grouped into three types of models and shown in Figure 2, models 1–26:

- Increasing Gaussian variance. Models 1–10 (Figure 2, top row) are realizations of the same Gaussian random field defined by an exponential covariance model with horizontal range of  $r_{\text{hor}} = 15$  m, vertical range of  $r_{\text{ver}} = 2$  m, mean velocity of 0.14 m/ns, and with an increasing standard deviation of  $\sigma = [0, 0.0040, 0.0057, 0.0069, 0.0080, 0.0089, 0.0098, 0.0106, 0.0113, 0.0120]$  m/ns. Thus, the variability is increasing, and the spatial correlation lengths are kept constant.
- Increasing Gaussian spatial correlation length. Models 11–20 (Figure 2, middle row) are realizations of the same Gaussian random field defined by an exponential covariance model with mean velocity of 0.14 m/ns, constant standard deviation of  $\sigma = 0.0120$  m/ns, but with increasing

horizontal range of  $r_{\text{hor}} = [0.01, 1.12, 2.23, 3.34, 4.45, 5.56, 6.67, 7.78, 8.89, 10.0]$  m, assuming a constant anisotropy of factor of  $r_{\text{hor}}/r_{\text{ver}} = 5$ . Thus, the variance is the same, but the spatial correlation lengths are increasing.

- Increasing multiple-point variability. Models 21–26 (Figure 2, bottom row) are realizations of a multiple-point-based statistical model inferred from the binary training image of Figure 3 (taken from [Strebelle, 2000](#)). The actual realization is generated using SNESIM ([Strebelle, 2002](#)). The location of the channel structure is the same for all six models, but the velocity within the channel is decreasing as  $v_{\text{channel}} = [0.13, 0.12, 0.11, 0.10, 0.09, 0.08]$  m/ns, while the velocity in the region outside the channel increases as  $v_{\text{outside}} = [0.15, 0.16, 0.17, 0.18, 0.19, 0.20]$  m/ns. Thus, the relative velocity contrast at the channel edges is increasing.

For all of these 26 models, a “reference” traveltime data set,  $\mathbf{t}_{\text{exact}}$ , is computed by assuming that the “exact” forward model  $\mathbf{g}_{\text{exact}}$  is given by the  $\mathbf{g}_{\text{FW}}$  forward model described above. As a reference geometry, we consider 11 source and 41 receiver locations in the left and right boreholes, respectively, as shown in Figure 2, model 1. Only traveltimes related to raypaths dipping less than  $45^\circ$  are used, to mimic a real situation ([Looms et al., 2008b](#)). Each traveltime data set consists of 331 traveltimes. In addition, the traveltimes obtained using the different approximate forward models related to the  $\mathbf{g}_{\text{SR}}$ ,  $\mathbf{g}_{\text{BR}}$ ,  $\mathbf{g}_{\text{SFR}}$ ,  $\mathbf{g}_{\text{BFR}}$ , and  $\mathbf{g}_{\text{Born}}$ , will be computed and referred to as  $\mathbf{t}_{\text{SR}}$ ,  $\mathbf{t}_{\text{BR}}$ ,  $\mathbf{t}_{\text{FR}}$ ,  $\mathbf{t}_{\text{BFR}}$ , and  $\mathbf{t}_{\text{Born}}$ , respectively. This will allow quantification of the modeling error as the difference in traveltime due to using the approximate forward models and the exact forward model.

Figure 4 shows corresponding modeling error in the form of the mean and standard deviation of the error introduced by a specific choice of kernel for the 331 traveltime estimates for each of the 26 considered models. The actual data behind Figure 4 are shown in

Tables 1–3. The mean can be thought of as a bias, the average mean difference in traveltime, and the standard deviation reflects the magnitude of the modeling error.

Figure 4 (top row) illustrates clearly that for a homogeneous model (model 1), all kernels perform equally well. This is no surprise because the kernels are normalized such that the integral of each kernel is the ray length between the source and the receiver. More importantly, it clearly shows that as the subsurface variability increases, the absolute value of the mean and the standard deviation

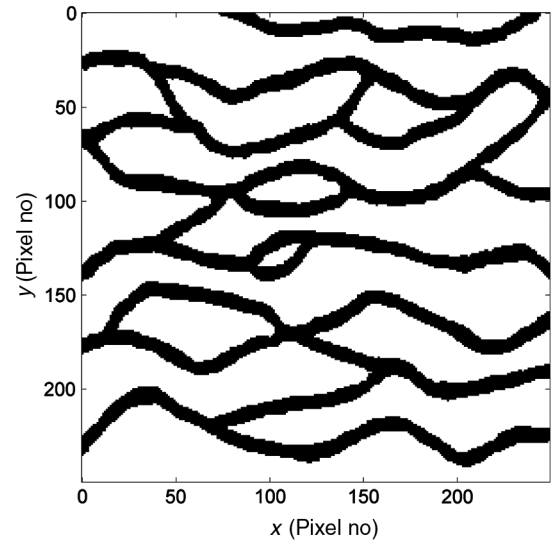


Figure 3. Bimodal training images, used to generate the realizations shown in Figure 2, models 21–26. From [Strebelle \(2000\)](#).

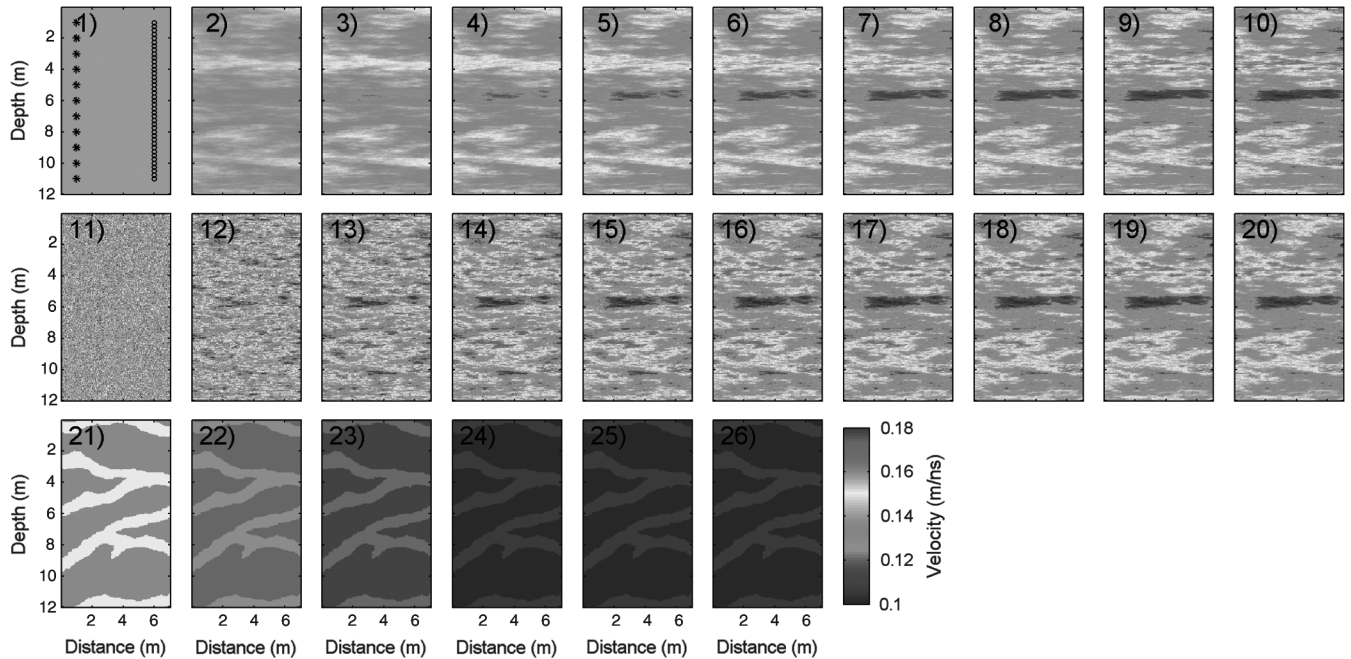


Figure 2. Twenty-six reference models. (Top row) Models 1–10 of type A (increasing variance). (Middle row) Models 11–20 of type B (increasing correlation length and constant variance). (Bottom row) Models 21–26 of type C (increasing velocity contrast for a binary velocity model). The location of receivers (stars, to the left) and sources (circles, to the right) is denoted on model 1. See the text for more details.

of the modeling error also increase using any of the approximations. A similar pattern can be seen for models 21–26, in which the velocity contrast in the binary velocity field is increasing. The absolute value of the mean and the standard deviation is significantly higher considering the discrete velocity models of models 21–26, than the models based on Gaussian statistics in models 1–10.

Figure 4 (middle row) shows the influence on the bias and variance of the modeling error for increasing spatial correlation lengths. The high-frequency approximations (ray sensitivities) are relatively more influenced by the changing correlation length than those based on Fresnel volumes. Notice that intermediate spatial wavelengths of the subsurface structures provide a larger modeling error than very high and low correlation lengths. The Fresnel volume sensitivities are less sensitive to the subsurface variability than the ray sensitivities because these sensitivities integrate the velocities of a larger volume of the subsurface.

Using the  $\mathbf{g}_{\text{BR}}$  forward model (high-frequency approximation to the wave equation) provides the fastest traveltime estimates because

the traveltime is computed along the fastest raypath between the transmitter and receiver. This is seen as a negative bias in the modeling error. The magnitude of the bias changes from zero (when the velocity model is constant) to more than  $-2$  ns for the velocity model with channels and maximum velocity contrast, model 26, Figure 4.

Modeling errors related to Gaussian variability of the velocity structures (models 1–20) provide a mean error up to 0.6 ns and a standard deviation of up to 0.6 ns. On the other hand, the modeling errors related to variability in the binary velocity structures (models 21–26) provide a mean error up to 12 ns and a standard deviation of up to 8 ns.

Measurement errors in a typical GPR cross-borehole data set are  $\sim 0.2$ – $0.8$  ns. Thus, for Gaussian-type velocity models, the modeling error can be up to the order of the uncertainty of the noise model. For the binary velocity distribution, the modeling error has the potential to be more than an order of magnitude larger than the noise model.

Figure 4 clearly demonstrates that the modeling error inherent in cross-borehole tomography can be significant, as compared to the measurement uncertainties. Furthermore, Figure 4 shows that the magnitude of the modeling error is closely linked to subsurface variability. As the subsurface variability increases, so does the magnitude of the modeling error.

In the following, we will propose a method to quantify the modeling error in a probabilistic framework. We will generate a sample of a probability distribution describing the modeling error and suggest how to account for such a modeling error when solving inverse problems.

## MODELING ERROR AND INVERSE PROBLEM THEORY

The methodology we will propose originates from a probabilistic formulation of inverse problem theory (see Tarantola, 2005). The solution to a probabilistically formulated inverse problem is a probability density obtained by combining all available states of information. Prior information on the model and data parameters, obtained independently, can in general be represented by the a priori probability distribution  $\rho(\mathbf{d}, \mathbf{m})$ . Information about the relation between data and model parameters can in general be represented by the theoretical probability density  $\Theta(\mathbf{d}, \mathbf{m})$ . The combined information is given by the joint posterior probability defined in the joint data and model space manifold  $\mathcal{D} \times \mathcal{M}$ :

$$\sigma(\mathbf{d}, \mathbf{m}) = k \frac{\rho(\mathbf{d}, \mathbf{m})\Theta(\mathbf{d}, \mathbf{m})}{\mu(\mathbf{d}, \mathbf{m})}. \quad (2)$$

The presence of  $\mu(\mathbf{d}, \mathbf{m})$  in equation 2 represents the homogeneous state of information that ensures that the parameterization is invariant to

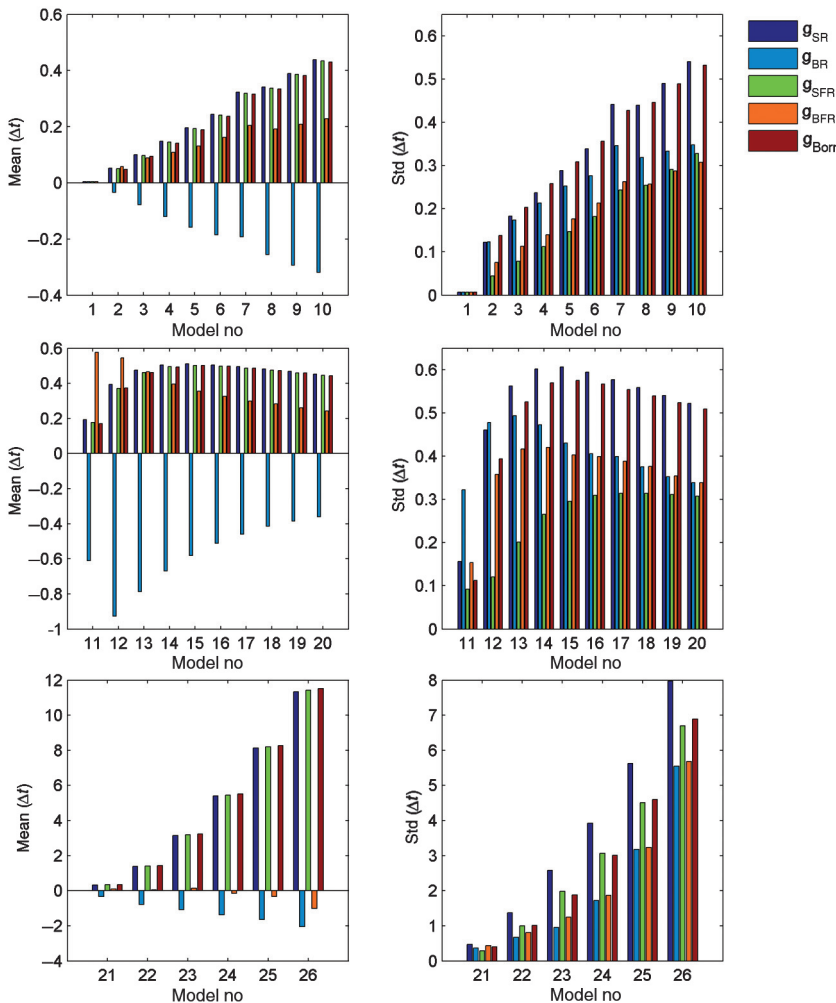


Figure 4. Observed mean (left) and standard deviation (right) of difference in traveltimes between the traveltimes obtained using the exact forward model  $\mathbf{t}_{\text{SR}}$  and approximate traveltime estimates  $\mathbf{t}_{\text{SR}}$ ,  $\mathbf{t}_{\text{BR}}$ ,  $\mathbf{t}_{\text{FR}}$ ,  $\mathbf{t}_{\text{BFR}}$ , and  $\mathbf{t}_{\text{Born}}$ , for velocity models of types A (models 1–10, top row), B (models 11–20, middle row), and C (models 21–26, bottom row).

changes in the coordinate system. Equation 2 is the most general way to define the solution to an inverse problem in a probabilistic framework (see, e.g., Tarantola and Valette, 1982b; Mosegaard and Tarantola, 2002). One may wish to infer information about the model parameters  $\mathbf{m}$  through the posterior marginal distribution of  $\mathbf{m}$  as given by

$$\sigma_M(\mathbf{m}) = \int_{\mathcal{D}} d\mathbf{d} \sigma(\mathbf{d}, \mathbf{m}). \quad (3)$$

If it is assumed that  $\rho_M(\mathbf{d})$  is obtained independently of  $\rho_M(\mathbf{m})$  such that  $\rho(\mathbf{d}, \mathbf{m}) = \rho_D(\mathbf{d})\rho_M(\mathbf{m})$  and one assumes the theoretical probability density can be given by

**Table 1. Observed mean (top) and standard deviation (bottom) of difference in traveltime between the traveltime picked from finite-difference data and the traveltime computed from sensitivity kernel models of type A.**

Model	1	2	3	4	5	6	7	8	9	10
$\mathbf{g}_{\text{SR}}$	0.0039	0.0522	0.0995	0.1474	0.1956	0.2439	0.3221	0.3408	0.3894	0.4385
$\mathbf{g}_{\text{BR}}$	0.0039	0.0240	-0.0778	-0.1192	-0.1569	-0.1843	-0.1919	-0.2550	-0.2927	-0.3192
$\mathbf{g}_{\text{SFR}}$	0.0039	0.0509	0.0976	0.1450	0.1929	0.2408	0.3186	0.3370	0.3853	0.4341
$\mathbf{g}_{\text{BFR}}$	0.0039	0.1157	0.0891	0.1083	0.1311	0.1623	0.2048	0.1926	0.2088	0.2289
$\mathbf{g}_{\text{SR}}$	0.0071	0.1220	0.1829	0.2366	0.2881	0.3389	0.4417	0.4399	0.4899	0.5406
$\mathbf{g}_{\text{BR}}$	0.0071	0.1229	0.1733	0.2132	0.2525	0.2763	0.3458	0.3182	0.3335	0.3476
$\mathbf{g}_{\text{SFR}}$	0.0071	0.0448	0.0783	0.1121	0.1472	0.1823	0.2437	0.2545	0.2907	0.3279
$\mathbf{g}_{\text{BFR}}$	0.0071	0.0755	0.1130	0.1395	0.1767	0.2131	0.2629	0.2571	0.2871	0.3074

**Table 2. Observed mean (top) and standard deviation (bottom) of difference in traveltime between the traveltime picked from finite-difference data and the traveltime computed from sensitivity kernel models of type B.**

Model	11	12	13	14	15	16	17	18	19	20
$\mathbf{g}_{\text{SR}}$	0.1928	0.3939	0.4749	0.5047	0.5103	0.5051	0.4940	0.4813	0.4671	0.4525
$\mathbf{g}_{\text{BR}}$	-0.6096	-0.9277	-0.7865	-0.6687	-0.5810	-0.5119	-0.4584	-0.4153	-0.3854	-0.3602
$\mathbf{g}_{\text{SFR}}$	0.1760	0.3718	0.4613	0.4949	0.5020	0.4973	0.4865	0.4738	0.4597	0.4451
$\mathbf{g}_{\text{BFR}}$	0.5761	0.5449	0.4647	0.3947	0.3560	0.3266	0.2977	0.2825	0.2611	0.2414
$\mathbf{g}_{\text{SR}}$	0.1561	0.4604	0.5625	0.6023	0.6063	0.5946	0.5775	0.5589	0.5404	0.5224
$\mathbf{g}_{\text{BR}}$	0.3225	0.4778	0.4940	0.4724	0.4302	0.4060	0.3994	0.3753	0.3523	0.3389
$\mathbf{g}_{\text{SFR}}$	0.0926	0.1204	0.2014	0.2656	0.2958	0.3092	0.3138	0.3143	0.3115	0.3075
$\mathbf{g}_{\text{BFR}}$	0.1535	0.3583	0.4164	0.4201	0.4031	0.3990	0.3885	0.3765	0.3541	0.3390

**Table 3. Observed mean (top) and standard deviation (bottom) of difference in traveltime between the traveltime picked from finite-difference data and the traveltime computed from sensitivity kernel models of type C.**

Model	21	22	23	24	25	26
$\mathbf{g}_{\text{SR}}$	0.3291	1.3800	3.1480	5.3911	8.1279	11.3291
$\mathbf{g}_{\text{BR}}$	-0.3353	-0.7892	-1.0850	-1.3806	-1.7341	-2.3468
$\mathbf{g}_{\text{SFR}}$	0.3417	1.4055	3.1874	5.4457	8.1997	11.4211
$\mathbf{g}_{\text{BFR}}$	0.1025	0.0566	0.1464	-0.1471	-0.3225	-1.0073
$\mathbf{g}_{\text{SR}}$	0.4675	1.3678	2.5801	3.9235	5.6202	7.9703
$\mathbf{g}_{\text{BR}}$	0.3663	0.6693	0.9500	1.7261	3.2388	5.8750
$\mathbf{g}_{\text{SFR}}$	0.2856	1.0039	1.9777	3.0595	4.5127	6.6988
$\mathbf{g}_{\text{BFR}}$	0.4330	0.8110	1.2512	1.8678	3.2283	5.7053



$$\Theta(\mathbf{d}, \mathbf{m}) = \theta(\mathbf{d}|\mathbf{m})\mu_M(\mathbf{m}), \quad (4)$$

where  $\mu_M(\mathbf{m})$  is the marginal probability density  $\mu_M(\mathbf{m}) = \int d\mathbf{m} \mu(\mathbf{d}, \mathbf{m})$ , then the solution to the inverse problem, of inferring information about the model parameters through the marginal a posteriori probability  $\sigma_M(\mathbf{m})$ , is given by

$$\sigma_M(\mathbf{m}) = k\rho_M(\mathbf{m})L(\mathbf{m}), \quad (5)$$

where  $k$  is a normalization constant and the likelihood function given by

$$L(\mathbf{m}) = \int_{\mathcal{D}} d\mathbf{d} \frac{\rho_D(g(\mathbf{m}))\theta(\mathbf{d}|\mathbf{m})}{\mu_D(\mathbf{d})}, \quad (6)$$

where  $\rho_D(\mathbf{d})$  describes *measurement uncertainties*, typically related to the instrument recording the data,  $\theta(\mathbf{d}|\mathbf{m})$  is a probabilistic formulation of the forward modeling that describes the probability of a set of calculated data given a model  $\mathbf{m}$ . This probability distribution is based on a probability distribution  $\Lambda(\mathbf{d}|\mathbf{m})$  that describes the uncertainty related to the forward calculation (i.e., a set of calculated data) and a (possibly uncertain) forward relation  $\mathbf{d} = \mathbf{g}(\mathbf{m})$ . In the following, we will refer to the probability distribution  $\Lambda(\mathbf{d}|\mathbf{m})$  as the *modeling error*. Based on modeling error, the probabilistic description of the forward modeling is given as  $\theta(\mathbf{d}|\mathbf{m}) = \Lambda(\mathbf{d} - \mathbf{g}(\mathbf{m}))$ .

Figure 5 provides a graphic illustration of  $\theta(\mathbf{d}|\mathbf{m})$  and  $\Lambda(\mathbf{d}|\mathbf{m})$  for a simple 1D inverse problem. In Figure 5a, the red line reflects an error-free theoretical relation  $\mathbf{d} = \mathbf{g}(\mathbf{m})$ , which implies that no

modeling error is present as shown in Figure 5e. In this case, the modeling error is expressed by a delta function,  $\Lambda(\mathbf{d}|\mathbf{m}) = \delta(\mathbf{d})$ . Figure 5b reflects a case in which modeling error is present in the forward modeling. Figure 5f shows the corresponding probability distribution reflecting the modeling errors  $\Lambda(\mathbf{d}|\mathbf{m}) = \theta(\mathbf{d} + \mathbf{g}(\mathbf{m})|\mathbf{m})$ . Note that, in this general case, the modeling error is described by a non-Gaussian distribution that is not centered around the theoretical relation  $\mathbf{d} = \mathbf{g}(\mathbf{m})$ , except at  $\mathbf{m} = 0$  where there is no modeling error.

For the remainder of the text, we will consider inverse problems whose solution is given by equation 5 and we will assume that the homogeneous probability density can be approximated by a constant, such that  $\mu_D(\mathbf{d}) = k$ . For more details on the homogeneous probability density function (pdf) (see, e.g., Mosegaard and Taran-tola, 2002).

### Accounting for the modeling error

In general, it is not straightforward to evaluate the modeling error  $\Lambda(\mathbf{d}|\mathbf{m})$  and hence perform the integration of equation 6.

At one extreme, the modeling error can be infinitely high, which means that any observed data will lead to the same constant likelihood  $L(\mathbf{m}) = k$ . In such a case, no information can be inferred from observed data and the a posteriori probability will be identical to the a priori probability  $\sigma_M(\mathbf{m}) = \rho_M(\mathbf{m})$ .

Another extreme is to completely ignore the modeling error. Disregarding the modeling error amounts to assuming a perfectly known noise-free relation between  $\mathbf{m}$  and  $\mathbf{d}$ . In such a case, the modeling error is described by a delta function, and we have

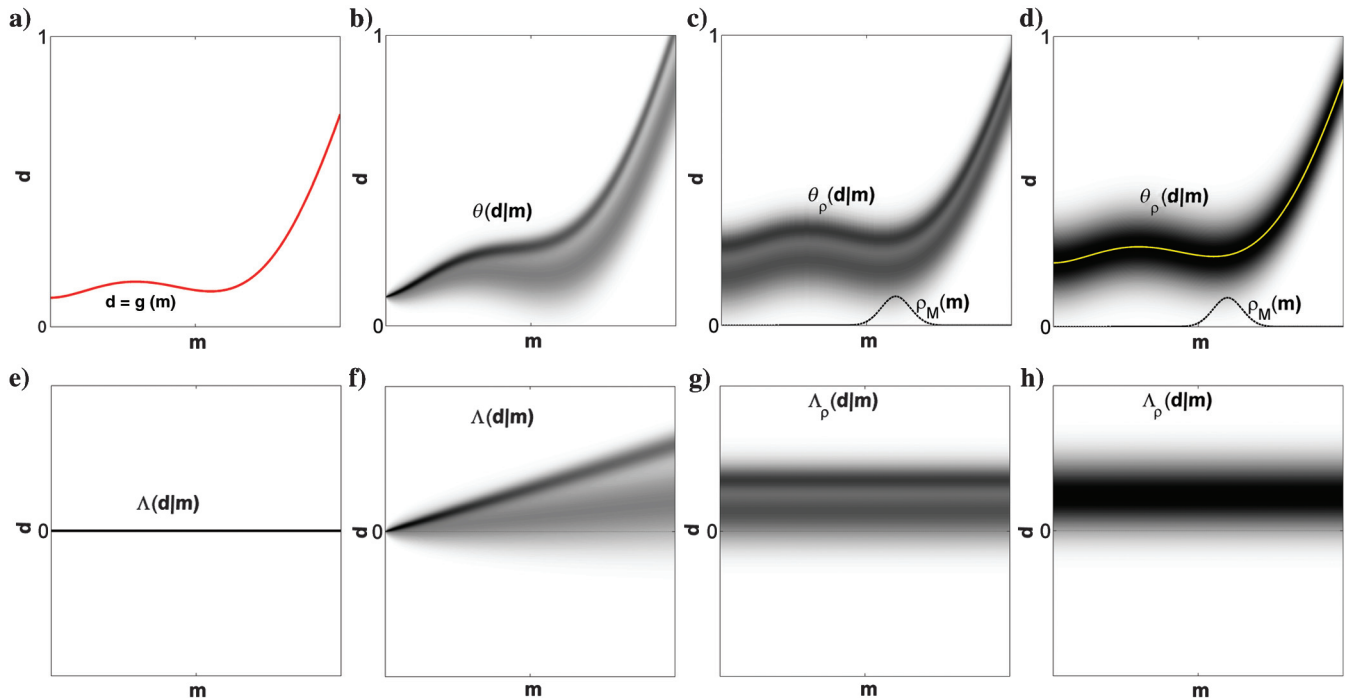


Figure 5. Schematic illustration of the modeling error. (a) In case of no theoretical errors,  $\theta(\mathbf{d}|\mathbf{m})$  is given by the red line  $\mathbf{d} = \mathbf{g}(\mathbf{m})$  and (e) the associated modeling error is given by  $\Lambda(\mathbf{d}|\mathbf{m}) = \delta(\mathbf{d})$ . (b) A general description of  $\theta(\mathbf{d}|\mathbf{m})$  and (f) the corresponding modeling error,  $\Lambda(\mathbf{d}|\mathbf{m})$ . (c)  $\theta_\rho(\mathbf{d}|\mathbf{m})$  in case of a (approximate) stationary modeling error  $\Lambda_\rho(\mathbf{d}|\mathbf{m})$  as seen in (g), obtained in the vicinity of  $\rho_M(\mathbf{m})$  (dotted line). (d)  $\theta_\rho(\mathbf{d}|\mathbf{m})$  in the case of an approximate stationary Gaussian modeling error  $\Lambda_\rho(\mathbf{d}|\mathbf{m}) = \mathcal{N}(\mathbf{d}_T, \mathbf{C}_T)$ , as seen in (h), obtained in the vicinity of  $\rho_M(\mathbf{m})$  (dotted line). The yellow curve reflects the maximum of the distribution  $\theta_\rho(\mathbf{d}|\mathbf{m})$ .



$\theta(\mathbf{d}|\mathbf{m}) = \delta(\mathbf{d} - \mathbf{g}(\mathbf{m}))$  (see equation 1.92 in Tarantola [2005] and Figure 5a), which reduces equation 6 to

$$L(\mathbf{m}) = \rho_D(\mathbf{g}(\mathbf{m})). \quad (7)$$

Evaluation of equation 7 only requires the choice of a probability model that describes the measurement uncertainty. If such measurement uncertainty can be described by a Gaussian model, with mean  $\mathbf{d}_d$  (which is often set to zero) and covariance  $\mathbf{C}_d$ , the likelihood can be expressed as

$$\begin{aligned} L(\mathbf{m}) &= \rho_D(\mathbf{m}) \\ &= k \exp\left(-\frac{1}{2}(\mathbf{g}(\mathbf{m}) - \mathbf{d}_{\text{obs}} - \mathbf{d}_d)^t \mathbf{C}_d^{-1} (\mathbf{g}(\mathbf{m}) - \mathbf{d}_{\text{obs}} - \mathbf{d}_d)\right). \end{aligned} \quad (8)$$

If significant modeling errors in reality exist, such as indicated by  $\theta(\mathbf{d}|\mathbf{m})$  in Figure 5b, and  $\Lambda(\mathbf{d}|\mathbf{m})$  in Figure 5f, then clearly ignoring such errors may lead to significant artifacts in describing the forward model and hence the solution to the inverse problem  $\sigma_M(\mathbf{m})$ . This will be further investigated and exemplified in a case study.

#### Accounting for Gaussian modeling errors

In the special case in which the modeling error can be described by a Gaussian probability density, with mean  $\mathbf{d}_T$  and covariance  $\mathbf{C}_T$ ,  $\mathcal{N}(\mathbf{d}_T, \mathbf{C}_T)$ , one can account for the measurement uncertainties and modeling error through addition of the covariance models describing measurement uncertainties  $\mathbf{C}_d$  and modeling errors  $\mathbf{C}_T$  (generalized after Tarantola [1986], p. 58):

$$\begin{aligned} L(\mathbf{m}) &= k \exp\left(-\frac{1}{2}(\mathbf{g}(\mathbf{m}) - \mathbf{d}_{\text{obs}} - \mathbf{d}_D)^t \mathbf{C}_D^{-1} (\mathbf{g}(\mathbf{m}) \right. \\ &\quad \left. - \mathbf{d}_{\text{obs}} - \mathbf{d}_D)\right), \end{aligned} \quad (9)$$

where  $\mathbf{C}_D = \mathbf{C}_d + \mathbf{C}_T$  and  $\mathbf{d}_D = \mathbf{d}_d + \mathbf{d}_T$ . Equation 9 is thus identical to equation 8 except that  $\mathbf{C}_d$  is replaced by the combined covariance model  $\mathbf{C}_D$  and a bias correction  $\mathbf{d}_D$  is introduced.

This means that one does thus not need to explicitly perform the integration of equation 6. In addition, equation 9 is valid for linear and nonlinear inverse problems, as long as the modeling error and the measurement uncertainty can be described by Gaussian statistics, in the form of  $\mathcal{N}(\mathbf{d}_T, \mathbf{C}_T)$  and  $\mathcal{N}(\mathbf{d}_d, \mathbf{C}_d)$ . If a valid Gaussian distributed modeling error can be established, then it can easily be used by many types of existing inversion algorithms that can account for Gaussian measurement uncertainties.

In the following, we will derive and test methods to approximate the modeling error as an additive correlated Gaussian error. First, we suggest methods for generating a sample of the modeling error, that may or may not be Gaussian distributed. Then, we suggest how to infer a Gaussian distributed modeling error based on such a sample from the probability distribution describing the modeling error.

## Quantifying modeling error

### Generating a sample of the modeling error

Let  $\mathbf{M} = [\mathbf{m}_1', \mathbf{m}_2', \dots, \mathbf{m}_N']$  represent a sample, in the form of  $N$  realizations from an a priori probability distribution  $\rho_M(\mathbf{m})$ . Consider two (linear or nonlinear) forward models in the form of an exact  $\mathbf{g}_{\text{ex}}$  and an approximate  $\mathbf{g}_{\text{app}}$  forward model. Let  $\mathbf{M}$  represent  $N$  realizations of the prior model  $\rho_M(\mathbf{m})$ . The corresponding data  $\mathbf{D}_{\text{ex}}$  and  $\mathbf{D}_{\text{app}}$  related to the exact and approximate forward models, respectively, can be computed as  $\mathbf{D}_{\text{ex}} = \mathbf{g}_{\text{ex}}(\mathbf{m})$ ,  $\mathbf{D}_{\text{app}} = \mathbf{g}_{\text{app}}(\mathbf{m})$ .

Each column in  $\mathbf{D}_{\text{ex}}$  and  $\mathbf{D}_{\text{app}}$  contains the data, e.g.,  $\mathbf{d}_{\text{ex},i}$  or  $\mathbf{d}_{\text{app},i}$ , related to the  $i$ th realization of the a priori probability density  $\rho_M(\mathbf{m})$ . A sample, in the form of  $N$  realizations, from the (unknown) probability distribution that describes the modeling error  $\Lambda_\rho(\mathbf{d}|\mathbf{m})$  is now available as

$$\mathbf{D}_\Lambda = \mathbf{D}_{\text{ex}} - \mathbf{D}_{\text{app}}. \quad (10)$$

We refer to the obtained sample of the modeling error as a sample from  $\Lambda_\rho(\mathbf{d}|\mathbf{m})$  as it is related to the subset of all models defined by the prior model,  $\rho_M(\mathbf{m})$ . To generate a sample of the full modeling error  $\Lambda(\mathbf{d}|\mathbf{m})$  one would need to consider any model, independent of its a priori likelihood. This will in practice not be possible. We therefore choose to assume stationarity of the modeling error, around a priori acceptable models, such that the sample of the modeling error obtained is valid in the vicinity of a priori acceptable models.

Figure 5g illustrates such a stationary modeling error  $\Lambda_\rho(\mathbf{d}|\mathbf{m})$  that would be obtained from the general modeling error  $\Lambda(\mathbf{d}|\mathbf{m})$  shown in Figure 5f, using the a priori distribution  $\rho_M(\mathbf{m})$  shown on the  $x$ -axis in Figure 5c, illustrating  $\theta_\rho(\mathbf{d}|\mathbf{m})$ . In the vicinity of the prior model,  $\Lambda_\rho(\mathbf{d}|\mathbf{m})$  is a good approximation of the general modeling error  $\Lambda(\mathbf{d}|\mathbf{m})$ . For models with a small a priori likelihood, it is also seen that  $\Lambda_\rho(\mathbf{d}|\mathbf{m})$  does not reflect  $\Lambda(\mathbf{d}|\mathbf{m})$  very well. Note how the shape of the distribution of the modeling error is the same for all values of  $\mathbf{m}$ , due to the assumption of stationarity. It is shifted vertically by a constant bias relative to  $\mathbf{d} = \mathbf{g}(\mathbf{m})$ . For well-posed inverse problems, one should not need consider a priori models with very low a priori likelihood, and hence  $\Lambda_\rho(\mathbf{d}|\mathbf{m})$  should be a reasonable approximation of  $\Lambda(\mathbf{d}|\mathbf{m})$ . If the general modeling error is in fact stationary as depicted in Figure 5g, then  $\Lambda_\rho(\mathbf{d}|\mathbf{m})$  will be identical to  $\Lambda(\mathbf{d}|\mathbf{m})$  as  $\Lambda(\mathbf{d}|\mathbf{m})$  will be the same for all  $\mathbf{m}$ .

### Approximate estimation of a Gaussian modeling error

In case the sample  $\mathbf{D}_\theta$  of modeling errors can be seen as a sample from a Gaussian probability density, we can readily estimate the mean value  $\mathbf{d}_{T_{\text{app}}}$  (the bias) and covariance matrix  $\mathbf{C}_{T_{\text{app}}}$  of this Gaussian distribution ( $\Lambda_\rho(\mathbf{d}|\mathbf{m}) = \mathcal{N}(\mathbf{d}_{T_{\text{app}}}, \mathbf{C}_{T_{\text{app}}})$ ) by the following equations:

$$\begin{aligned} \mathbf{d}_{T_{\text{app}}} &= [\mathbf{d}_{T_{\text{app}}}^1, \mathbf{d}_{T_{\text{app}}}^2, \dots, \mathbf{d}_{T_{\text{app}}}^N] \\ \text{where } \mathbf{d}_{T_{\text{app}}}^i &= \frac{1}{N} \sum_{j=1}^N (D_{\text{ex}}^{i,j} - D_{\text{app}}^{i,j}), \end{aligned} \quad (11)$$

$$\mathbf{C}_{T_{\text{app}}} = \frac{1}{N} \mathbf{D}_{\text{diff}} \mathbf{D}_{\text{diff}}' \quad \text{where} \quad \mathbf{D}_{\text{diff}} = [\mathbf{D}_{\text{ex}} - \mathbf{D}_{\text{app}} - \mathbf{D}_{T_{\text{app}}}], \quad (12)$$

and  $\mathbf{D}_{T_{\text{app}}} = [\mathbf{d}_{T_{\text{app}}}^1, \mathbf{d}_{T_{\text{app}}}^2, \dots, \mathbf{d}_{T_{\text{app}}}^N]$ .

Figure 5h illustrates such a Gaussian distributed modeling error  $\Lambda_\rho(\mathbf{d}|\mathbf{m}) = \mathcal{N}(\mathbf{d}_{T_{\text{app}}}, \mathbf{C}_{T_{\text{app}}})$  as inferred from  $\mathbf{D}_{\text{ex}}$  and  $\mathbf{D}_{\text{app}}$  using equations 11–12, and Figure 5d the corresponding  $\theta_\rho(\mathbf{d}|\mathbf{m})$ . Again, due to the assumption of stationarity, as Figure 5c and 5g, the Gaussian distribution of the modeling error is the same for all values of  $\mathbf{m}$ , and defined by the variance  $\mathbf{C}_T$  and shifted by a constant  $\mathbf{d}_T$  relative to  $\mathbf{d} = \mathbf{g}(\mathbf{m})$ .

Whether or not  $\mathcal{N}(\mathbf{d}_{T_{\text{app}}}, \mathbf{C}_{T_{\text{app}}})$  is a good description of the actual modeling error  $\Lambda_\rho(\mathbf{d}|\mathbf{m})$  naturally depends on whether the Gaussian assumption is valid.

#### Exact calculation of a Gaussian modeling error in the linear Gaussian case

For linear Gaussian inverse problems, the forward problem can be quantified by the forward operator  $\mathbf{G}$ , such that  $\mathbf{d} = \mathbf{G}\mathbf{m}$ . In addition, the a priori information on the model parameters is described by the Gaussian model  $\mathcal{N}(\mathbf{m}_0, \mathbf{C}_M)$ . Using  $\mathbf{G}$ ,  $\mathbf{m}_0$ , and  $\mathbf{C}_M$ , one can map the a priori information about the model parameters into an a priori Gaussian distribution of data  $\mathcal{N}(\mathbf{d}_\rho, \mathbf{C}_\rho)$ , with

$$\mathbf{d}_\rho = \mathbf{G}\mathbf{m}_0, \quad (13)$$

$$\mathbf{C}_\rho = \mathbf{G}\mathbf{C}_M\mathbf{G}'. \quad (14)$$

The symbol  $\mathbf{C}_\rho$  should not be confused with data measurement uncertainties, but is a description of the covariance between data given the choice of the a priori covariance model and the forward operator.

Consider an ideal case in which an exact expression of the linear forward operator is given as  $\mathbf{G}_{\text{ex}}$ . Another approximate linear forward kernel is given by  $\mathbf{G}_{\text{app}}$ . The difference in data computed using these two kernels is given by

$$\begin{aligned} \mathbf{d}_{\text{ex}} &= \mathbf{G}_{\text{ex}}(\mathbf{m} - \mathbf{m}_0) + \mathbf{G}_{\text{ex}}\mathbf{m}_0 \\ \mathbf{d}_{\text{app}} &= \mathbf{G}_{\text{app}}(\mathbf{m} - \mathbf{m}_0) + \mathbf{G}_{\text{app}}\mathbf{m}_0 \\ &\Downarrow \\ \mathbf{d}_{\text{ex}} - \mathbf{d}_{\text{app}} - [\mathbf{G}_{\text{ex}} - \mathbf{G}_{\text{app}}]\mathbf{m}_0 &= [\mathbf{G}_{\text{ex}} - \mathbf{G}_{\text{app}}](\mathbf{m} - \mathbf{m}_0) \\ &\Downarrow \\ \mathbf{d}_{\text{ex}} - \mathbf{d}_{\text{app}} - \mathbf{d}_T &= [\mathbf{G}_{\text{ex}} - \mathbf{G}_{\text{app}}](\mathbf{m} - \mathbf{m}_0). \end{aligned} \quad (15)$$

Following equation 16, it is evident that the difference between the exact and approximate estimates of data  $\mathbf{d}_{\text{ex}} - \mathbf{d}_{\text{app}}$  is linearly related to  $\mathbf{m}$  through the difference between the linear kernels  $\mathbf{G}_{\text{ex}} - \mathbf{G}_{\text{app}}$ . Following equations 13–14, a complete description of the Gaussian modeling error  $\mathcal{N}(\mathbf{d}_T, \mathbf{C}_T)$  can now be given as

$$\mathbf{d}_T = [\mathbf{G}_{\text{ex}} - \mathbf{G}_{\text{app}}]\mathbf{m}_0, \quad (17)$$

$$\mathbf{C}_T = [\mathbf{G}_{\text{ex}} - \mathbf{G}_{\text{app}}]\mathbf{C}_M[\mathbf{G}_{\text{ex}} - \mathbf{G}_{\text{app}}]'. \quad (18)$$

$\mathcal{N}(\mathbf{d}_T, \mathbf{C}_T)$  provides an exact description of the modeling error  $\Lambda_\rho(\mathbf{d}|\mathbf{m})$  as a result of using the approximate (and linear) kernel  $\mathbf{G}_{\text{app}}$  given a known exact (and linear) kernel  $\mathbf{G}_{\text{ex}}$  and the a priori information given by  $\mathcal{N}(\mathbf{m}_0, \mathbf{C}_M)$ . This is an exact description of the modeling error approximated in equations 11–12.

#### Example of estimating a Gaussian modeling error

The exact method described above applies to linear problems with a Gaussian prior model and Gaussian data uncertainties. The approximate method is completely general and provides a sample of the modeling error. The larger the sample, the better statistics about the modeling error can be inferred. We will illustrate this for a linear case in which the exact method can serve as the benchmark. We consider two linear forward models defined previously where  $\mathbf{g}_{\text{SFR}}$  shall serve as the “true” forward model and  $\mathbf{g}_{\text{SR}}$  serve as the approximate forward model.

Figure 6a–6h shows the approximate estimates of the covariance of a Gaussian modeling error,  $\mathbf{C}_{T_{\text{app}}}$ , obtained using equation 12 and  $N = 5, 10, 25, 50, 100, 200, 400$ , and 800 realizations from the prior model. Figure 6i shows the exact estimate of  $\mathbf{C}_{T_{\text{app}}}$  computed using equation 18. The estimated mean of the Gaussian modeling error  $\mathbf{d}_{T_{\text{app}}}$  obtained using equation 11, is shown in Figure 6j for  $N = 10, 400$ , and 800 realizations and compared to the exact estimate of  $\mathbf{d}_{T_{\text{app}}}$  computed using equation 17. Figure 6 illustrates how the approximate estimates of  $\mathbf{d}_{T_{\text{app}}}$  and  $\mathbf{C}_{T_{\text{app}}}$  tend toward the exact estimates of  $\mathbf{d}_T$  and  $\mathbf{C}_T$  as the number of considered a priori realizations increases.

So, in this simple linear case, the general nonlinear approach works well and requires relatively few realizations of the modeling error. Below, we approach two more realistic problems with nonlinear forward models and, for one case, a non-Gaussian prior model where our general sampling method shows its potential.

## CASES

We have shown (see Figure 4), that the modeling error caused by the use of an imperfect forward model can be more than an order of magnitude larger than the measurement uncertainty. We have also proposed a method to quantify a Gaussian modeling error. Through two case studies, we will consider the effect of disregarding and accounting for the modeling error, respectively.

In the first case study, we investigate the effect of disregarding the modeling error when solving a nonlinear Gaussian inverse problem as a linear inverse Gaussian problem, using least-squares-type inversion.

In the second case study, we explore the use of a fast (and approximate) nonlinear forward model to solve a nonlinear inverse problem using a non-Gaussian a priori model. In this case, a sample from the a posteriori probability density is obtained using the extended Metropolis algorithm.

#### Non-Gaussian prior model

Figure 7a shows a reference velocity model, which is generated as a realization of Gaussian random field with mean value of 0.14 m/ns, and a covariance model with the distance decay known as “spherical” within geostatistics, with a horizontal range of 10 m and a vertical range of 2 m. From the reference velocity model, a

reference data set is computed using the forward-type model based on finite-difference waveform modeling  $\mathbf{g}_{FW}$ . Waveform data, from where the traveltime data are picked, are obtained using a Ricker wavelet with peak frequency of 100 MHz. The recording geometry is the same as that used in Figure 2, which leads to 331 traveltimes. In addition, three levels of measurement uncertainties  $\mathbf{C}_d$  have been added as uncorrelated Gaussian noise with a standard deviation of 0 (noise-free data), 0.1, and 0.4 ns, respectively (see Figure 7b). For the remainder of this case study, we will consider uncorrelated measurement uncertainty only, and we will refer to a model of uncorrelated measurement uncertainty with a standard deviation of, e.g., 0.1 ns as  $\mathbf{C}_d = \mathbf{I}0.1^2$ . We thus consider the three different data sets as observed data, contaminated with different measurement noise.

The true forward problem is nonlinear. We will invert these traveltime data using an approximate linear straight-ray forward model  $\mathbf{g}_{SR}$  with a known Gaussian a priori model  $\mathcal{N}(\mathbf{m}, \mathbf{C}_M)$  as used to generate the reference model. As we make use of an approximation to describe the forward problem, a modeling error will be introduced. Treated as a linear Gaussian inverse problem, the solution to the inverse problem is described by the Gaussian a posteriori probability density that can be completely characterized by the a posteriori mean  $\tilde{\mathbf{m}}$  and covariance model  $\tilde{\mathbf{C}}_M$ ,  $\mathcal{N}(\tilde{\mathbf{m}}, \tilde{\mathbf{C}}_M)$  (see, e.g., Tarantola and Valette, 1982):

$$\tilde{\mathbf{m}} = \mathbf{m}_0 + \mathbf{C}_M \mathbf{G}^t (\mathbf{G} \mathbf{C}_M \mathbf{G}^t + \mathbf{C}_D)^{-1} \times (\mathbf{d}_0 - \mathbf{G} \mathbf{m}_0), \quad (19)$$

$$\tilde{\mathbf{C}}_M = \mathbf{C}_M - \mathbf{C}_M \mathbf{G}^t (\mathbf{G} \mathbf{C}_M \mathbf{G}^t + \mathbf{C}_D)^{-1} \mathbf{G} \mathbf{C}_M. \quad (20)$$

The a posteriori mean model  $\tilde{\mathbf{m}}$  is also the model with maximum a posteriori probability.

#### Inversion with no modeling error

Initially, we will disregard the modeling error ( $\mathbf{d}_T = 0$  and  $\mathbf{C}_T = 0$ ) such that  $\mathbf{C}_D = \mathbf{C}_d$ . That is, we properly account for the three cases of measurement errors but ignore the modeling error. The linear inverse problem is solved given the three data sets using equations 19–20. Figure 8a–8c shows the corresponding maximum a posteriori models,  $\tilde{\mathbf{m}}$ , considering  $\mathbf{C}_d = \mathbf{I}0$  ns<sup>2</sup>,  $\mathbf{I}0.1$  ns<sup>2</sup>, and  $\mathbf{I}0.4$  ns<sup>2</sup>, respectively. Figure 8d–8f shows three realizations from the corresponding a posteriori probability density.

Using noise-free data, the inversion result is severely affected by disregarding the modeling error, Figure 8a and 8d. Within the area of high ray coverage (in the middle of the model), the maximum a posteriori model and realizations from the a posteriori probability density show abnormal velocity variation, in terms of magnitude

and spatial variability. This is a result of fitting the modeling error, which acts as noise, and is not accounted for. To fit the data according to the assumption of noise-free data, extreme velocity variations are introduced into the solution space. In addition, comparing the realizations from the a posteriori probability density, many of the structures, which are artifacts and a result of fitting modeling noise, seems to be well resolved. A feature is well resolved if the same feature appears in many realizations of the a posteriori probability density. Thus, ignoring modeling error in this case introduces artifacts, which appear well resolved, into the inversion results.

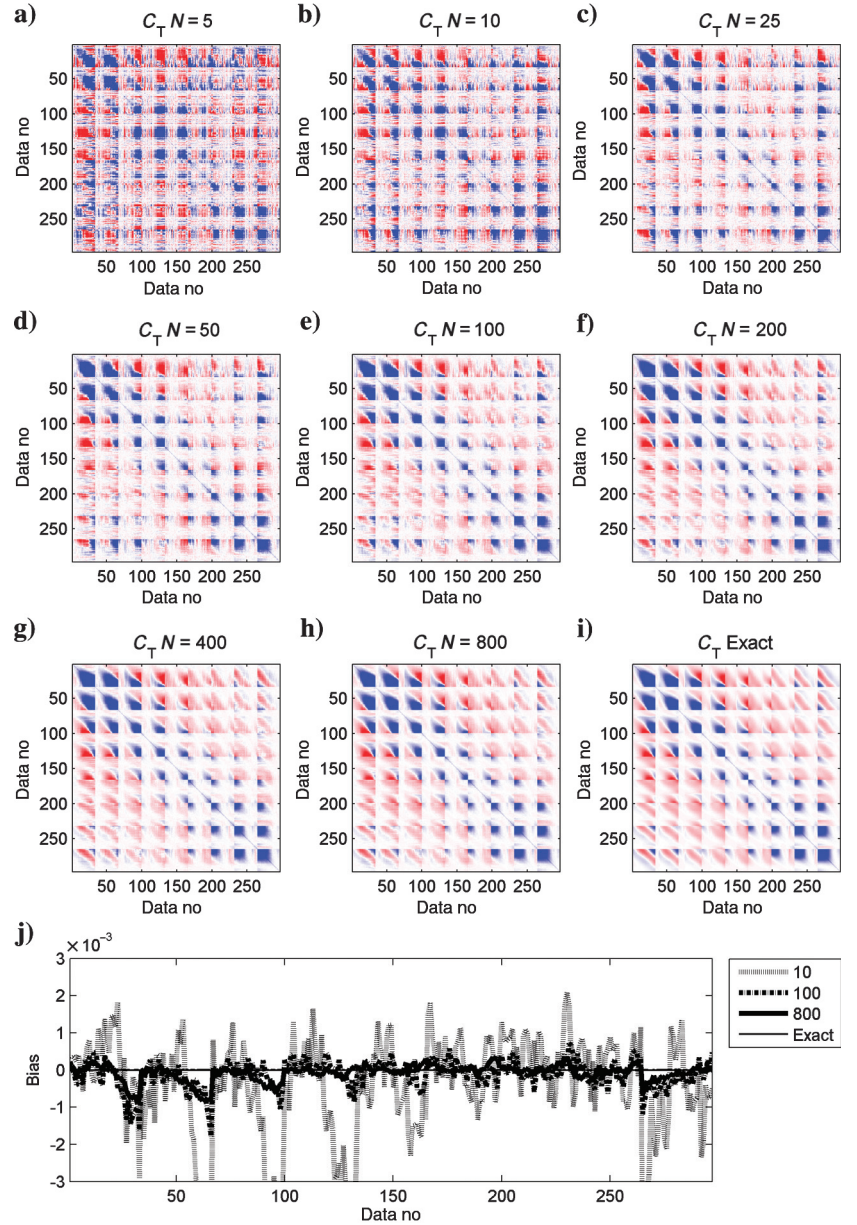


Figure 6. (a–h) Approximate estimate of  $\mathbf{C}_{T,app}$ , obtained using equation 12 for  $N = 5, 10, 25, 50, 100, 200, 400$ , and  $800$  realizations from the a priori pdf. (i) Exact computation of  $\mathbf{C}_T$  using equation 18. (j) Comparison of  $\mathbf{d}_{T,app}$ , obtained using equation 11 for  $N = 10, 100$ , and  $800$  realizations from the a priori pdf, to computing the exact  $\mathbf{d}_T$  using equation 17.



It is of course quite extreme to invert data without assuming any data noise at all. Considering more realistic measurement uncertainties of  $\mathbf{C}_d = \mathbf{I}0.1^2$ , the problems referred to above persist. The maximum a posteriori model in Figure 8b shows a low-velocity zone at  $x = 3.5$  m,  $y = 5.5$  m, surrounded by a high-velocity zone. This low-velocity zone seems to be well resolved because it can be identified on the three realizations from the a posteriori probability density, Figure 8e. However, compared to the reference model, Figure 7a, it is evident, that such a feature does not exist and is an effect of disregarding the modeling error. As the measurement uncertainty is increased, the effect of disregarding the modeling error is reduced (see Figure 8c and 8f). However, using  $\mathbf{C}_d = \mathbf{I}0.4$  ns<sup>2</sup>, a low-velocity zone still appears as a rather well resolved feature at  $x = 3.5$  m,  $y = 5.5$  m.

### Estimating Gaussian modeling error

To estimate the modeling error caused using the linear straight-ray approximation  $\mathbf{g}_{\text{SR}}$  600 realizations from the a priori Gaussian model  $\mathcal{N}(\mathbf{m}, \mathbf{C}_M)$  are generated. For each of the 600 models, the forward response is computed using the ideal forward method  $\mathbf{g}_{\text{FW}}$ , which provides  $D_{\text{ex}}$ . Likewise, the approximate linear forward model  $\mathbf{g}_{\text{SR}}$  is used to calculate the traveltimes in  $\mathbf{D}_{\text{app}}$ . Then, a Gaussian distributed modeling error  $\mathcal{N}(\mathbf{d}_{T_{\text{app}}}, \mathbf{C}_{T_{\text{app}}})$  can be established using equations 17–18. Figure 9a and 9b shows the estimated covariance  $\mathbf{C}_{T_{\text{app}}}$  and mean  $\mathbf{d}_{T_{\text{app}}}$  describing the inferred Gaussian distributed modeling error.

It has been assumed that the modeling error can be characterized by a Gaussian model, as given by equations 17–18. This may not always be a reasonable assumption. One simple way of checking whether the Gaussian model is reasonable is to compare actual realizations of the modeling error, as obtained in equation 10, to realizations of the inferred model of Gaussian modeling error. Figure 9c compares the actual modeling error (black curve) to one realization of the inferred Gaussian modeling error (red) for one realization of the prior model  $\mathcal{N}(\mathbf{m}, \mathbf{C}_M)$ . A visual comparison suggests that the estimated modeling variability resembles the expected/computed variability.

A more formal way to evaluate the Gaussian assumption is to, for example, evaluate if the  $\chi^2$ -distribution related to the computed modeling errors  $\mathbf{D}_\Lambda$  are consistent with the inferred Gaussian model describing modeling errors  $\mathcal{N}(\mathbf{d}_{T_{\text{app}}}, \mathbf{C}_{T_{\text{app}}})$ . If

$\mathbf{D}_\Lambda = [\mathbf{d}_{\Lambda,1}, \mathbf{d}_{\Lambda,2}, \dots, \mathbf{d}_{\Lambda,N}]$ , then  $N = 600$   $\chi^2$  values can be computed as

$$\chi_i^2 = (\mathbf{d}_{\Lambda,i} - \mathbf{d}_{T_{\text{app}}})^T \mathbf{C}_{T_{\text{app}}} (\mathbf{d}_{\Lambda,i} - \mathbf{d}_{T_{\text{app}}}). \quad (21)$$

The distribution of  $\chi^2$  values related to actual realizations generated from  $\mathcal{N}(\mathbf{m}, \mathbf{C}_M)$  will be distributed according to the chi-squared distribution with  $M$  degrees of freedom, where  $M = 331$  is the number of observed data (i.e., the length of  $\mathbf{d}_{T_{\text{app}}}$ ) (Tarantola, 2005). If  $\mathbf{D}_\Lambda$  is a sample from  $\mathcal{N}(\mathbf{m}, \mathbf{C}_M)$ , then the distribution of  $[\chi_1^2, \chi_2^2, \dots, \chi_N^2]$  should also be described by a chi-squared distribution with  $ND$  degrees of freedom. Figure 10 compares the distribution of  $\chi^2$  computed using equation 21, to the distribution of  $\chi^2$  obtained from an actual sample of  $\mathcal{N}(\mathbf{d}_{T_{\text{app}}}, \mathbf{C}_{T_{\text{app}}})$ . Note how both distributions are distributed around  $\chi^2 = 331$ , as should be expected. It is evident that there is no reason to assume that the Gaussian assumption is not valid.

### Inversion with modeling error

We now consider inversion of the traveltime data considering measurement uncertainty  $\mathbf{C}_d$  and, at the same time, Gaussian distributed modeling error,  $\mathcal{N}(\mathbf{d}_{T_{\text{app}}}, \mathbf{C}_{T_{\text{app}}})$ , using equations 19 and 20. Inversion results, in the form of the maximum a posteriori model, and three realizations from the corresponding a posteriori probability density, can be seen in Figure 11, which can be directly compared to the results of disregarding the modeling error in Figure 8.

From the realizations of the a posteriori probability density, Figure 11d–11f, it is evident that spatial variability is similar, using any of the three considered noise models, and is similar to the spatial variability of the reference model. The resolution decreases as the measurement uncertainties increase, as should be expected.

Figure 6 illustrates the accuracy of  $\mathcal{N}(\mathbf{d}_{T_{\text{app}}}, \mathbf{C}_{T_{\text{app}}})$  as a function of the sample size. Figure 12 illustrates the effect on the inversion result, in the form of the maximum a posteriori model,  $\tilde{\mathbf{m}}$ , related to the sample size used to compute  $\mathcal{N}(\mathbf{d}_{T_{\text{app}}}, \mathbf{C}_{T_{\text{app}}})$ . For the present case, a sample of at least  $N = 300$  realizations are needed to obtain stable inversion results.

It has been shown that it may be possible to solve a nonlinear inverse problem, with a Gaussian a priori model, using a fast approximate linear approach, in a probabilistic sound manner, if the modeling error caused by using an approximate forward model

can be described by a Gaussian model. The resulting solution to the inverse problem, in the form of the a posteriori probability density, is an approximation to the full inverse problem using the true forward model. This means that model parameters will be less resolved than if using the exact forward model. However, the result may be obtained using a much faster forward model. Moreover, in this particular case, the linear assumption provides a means to obtain a full characterization of the (Gaussian) posterior distribution using equations 19 and 20, instead of sampling the posterior distribution of a nonlinear inverse problem, which would have resulted in a huge number of forward calculations. An example where such a sample problem is considered is seen in the next section.

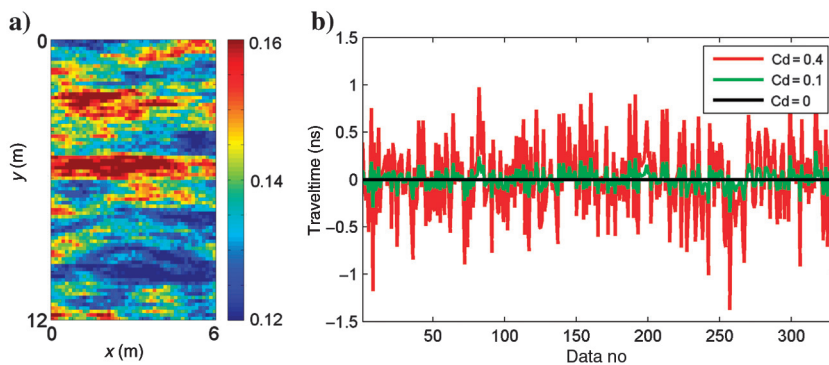


Figure 7. (a) Reference velocity model (m/ns). (b) Realizations of the Gaussian models for  $\mathbf{C}_d$ , with a standard deviation of 0, 0.1, and 0.4 ns.

### Case B: Accounting for modeling errors in non-linear cross-borehole tomography, with a non-Gaussian prior model

We will now consider a nonlinear inverse problem with a non-Gaussian a priori model. Figure 13 shows a clearly non-Gaussian

distributed reference model. It has been generated using the SNESIM algorithm using the image in Figure 3 as training image, [Strebel \(2000\)](#). The velocity within the thin channel structures is 0.18 m/ns and the velocity of the background material is 0.10 m/ns.

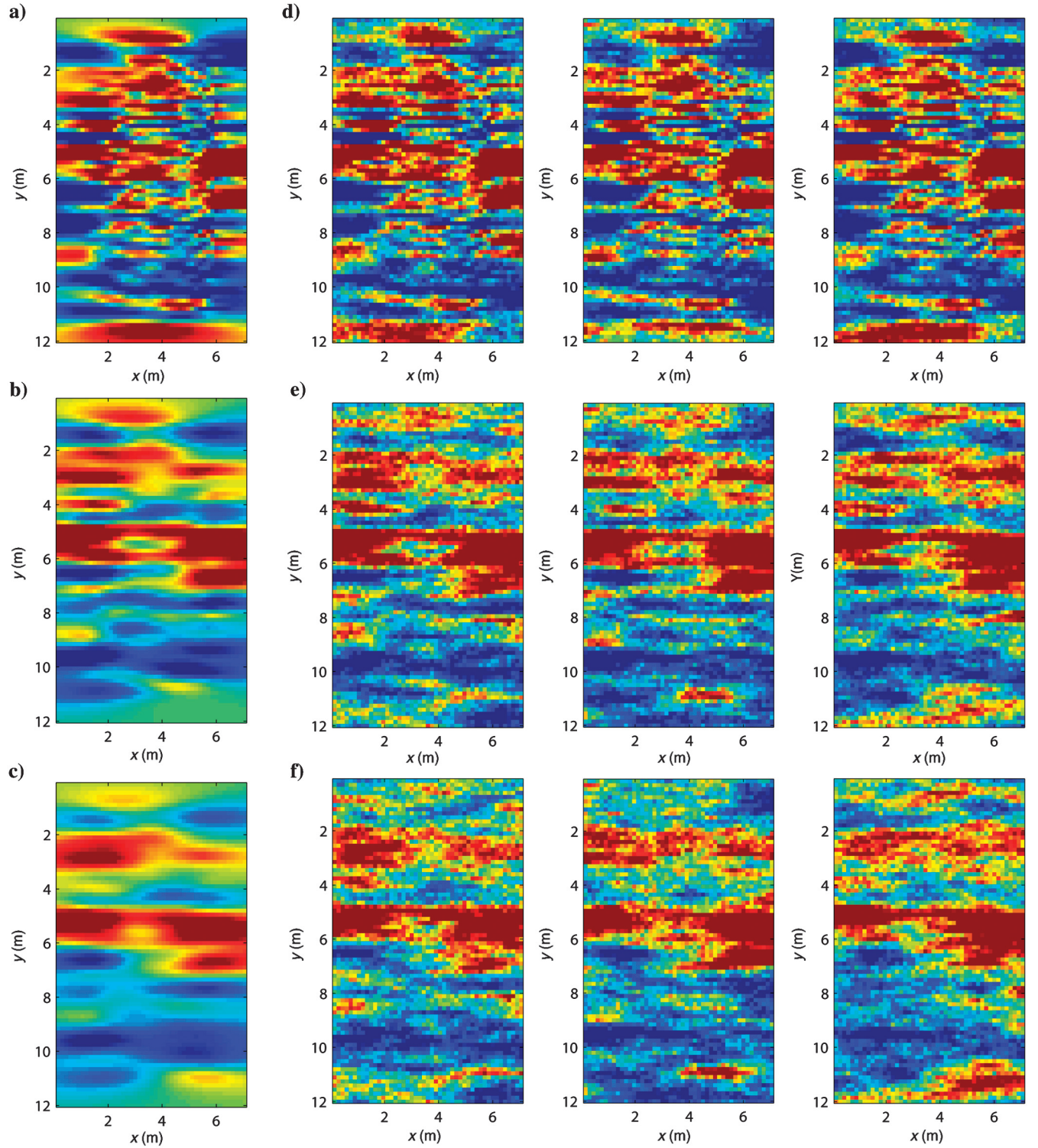


Figure 8. Inversion disregarding modeling error  $C_T = 0$ . (a-c) Maximum a posteriori model (least-squares mean estimate) for  $C_d = 10, 10.1^2$ , and  $10.4^2$ . (d-f) Three realizations from the a posteriori probability density considering  $C_d = 10, 10.1^2$ , and  $10.4^2$ .



As for case A, a reference data set is computed using the finite-difference-based  $\mathbf{g}_{\text{FW}}$ -type forward model. A realization of uncorrelated Gaussian noise with a variance of  $0.64 \text{ ns}^2$  is added to the reference data to resemble measurement uncertainty. We now wish to solve the inverse problem of inferring information about the velocity distribution, given “observed” traveltime data, and an a priori model described by the SNESIM algorithm using the training image in Figure 3. As a “nonperfect” forward model, we consider the eikonal solution to the wave equation, i.e., the  $\mathbf{g}_{\text{BR}}$  forward model. In this case, the forward problem is neither linear nor is the a priori model Gaussian. Therefore, we cannot make use of least-squares-based inversion as in case A. Instead, we resort to sampling techniques that, as opposed to providing an analytical description of the a posteriori probability density, generates a sample of the a posteriori probability density (Mosegaard and Sambridge, 2002). We specifically make use of a combination of the extended Metropolis sampler (Mosegaard and Tarantola, 1995), and the

sequential Gibbs sampler (Hansen et al., 2012), as made available in the SIPPI software package (Hansen et al., 2013a, 2013b).

#### Quantifying the modeling error

A Gaussian distributed modeling error is estimated in a similar manner as it was done for case A. Initially, 600 realizations are generated from the a priori model. For each of the 600 models, the forward response is computed using the ideal forward method, i.e., the  $\mathbf{g}_{\text{FW}}$ -type forward model, providing  $\mathbf{D}_{\text{ex}}$ . Likewise, the approximate forward model, based on the eikonal wave equation  $\mathbf{g}_{\text{BR}}$  is used to calculate the traveltimes in  $\mathbf{D}_{\text{app}}$ . Then, we estimate a Gaussian distributed modeling error  $\mathcal{N}(\mathbf{d}_{T_{\text{app}}}, \mathbf{C}_{T_{\text{app}}})$  using equations 11 and 12.

Figure 14a shows the estimated bias  $\mathbf{d}_{T_{\text{app}}}$  and Figure 14b the estimated covariance  $\mathbf{C}_{T_{\text{app}}}$ . Note that the bias itself fluctuates around 1.6 ns, which is significantly higher than the standard deviation of 0.8 ns describing the measurement uncertainty.

#### Sampling the a posteriori probability density

The extended Metropolis algorithm is run for 120,000 iterations in two cases: Initially assuming no modeling error ( $\mathbf{d}_T = 0$ ,  $\mathbf{C}_T = 0$ ), and subsequently we account for the modeling error making use of the estimated Gaussian distributed modeling error ( $\mathbf{d}_T = \mathbf{d}_{T_{\text{app}}}$ ,  $\mathbf{C}_T = \mathbf{C}_{T_{\text{app}}}$ ). The a priori model is the same in both cases.

Figure 15a shows a sample of the a priori probability density in the form of five different independent realizations from the a priori distribution. It is apparent that the a priori assumption reflects some high-velocity channellike structures traverse the model parameter space from left to right, with layers dipping up and down from left to right. In addition, some of the channel structures are connected. A corresponding sample from the posterior probability density in the case of disregarding modeling error is shown in Figure 15b and in the case of accounting for the modeling error in Figure 15c.

Disregarding modeling error (Figure 15b), reveals that most realizations within the posterior sample exhibit the same type of features at the same locations. There is relatively little variability between the individual realizations. The posterior sample shows

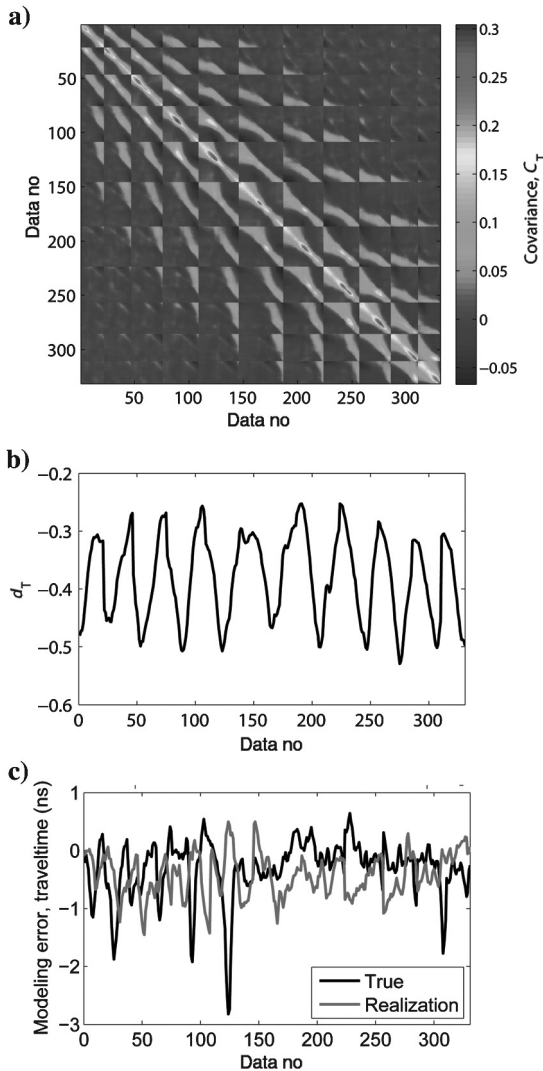


Figure 9. (a) Estimated Gaussian distributed modeling error,  $\mathbf{C}_{T_{\text{app}}}$ . (b) Estimated mean of modeling error  $\mathbf{d}_{T_{\text{app}}}$ . (c) Actual modeling error (dotted line) compared to one realization of the inferred modeling error (solid line) for one realization of  $\mathcal{N}(\mathbf{m}, \mathbf{C}_M)$ .

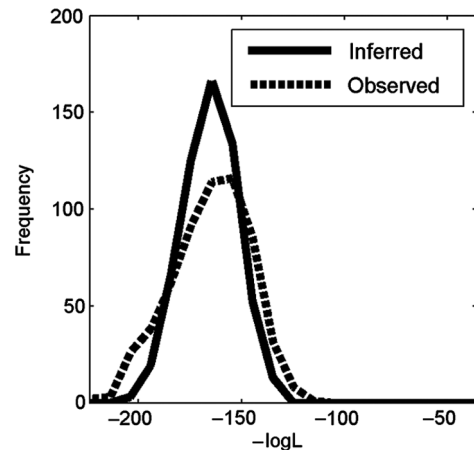


Figure 10. Distribution of  $\chi^2$  related to the inferred sample of modeling error ( $\mathbf{d}_A$ ) (dashed line) and actual realizations from  $\mathcal{N}(\mathbf{d}_{T_{\text{app}}}, \mathbf{C}_{T_{\text{app}}})$  (solid line).

considerably more variability between posterior realizations when accounting for the modeling error (Figure 15c). At a first glance, it appears as if the posterior sample obtained by disregarding the modeling error provides better resolution (i.e., less variability) of the

solution (i.e., posterior realizations), than the case of accounting for the modeling error.

This is further highlighted in Figure 16. Figure 16a and 16b shows the a posteriori probability of locating a channel in the

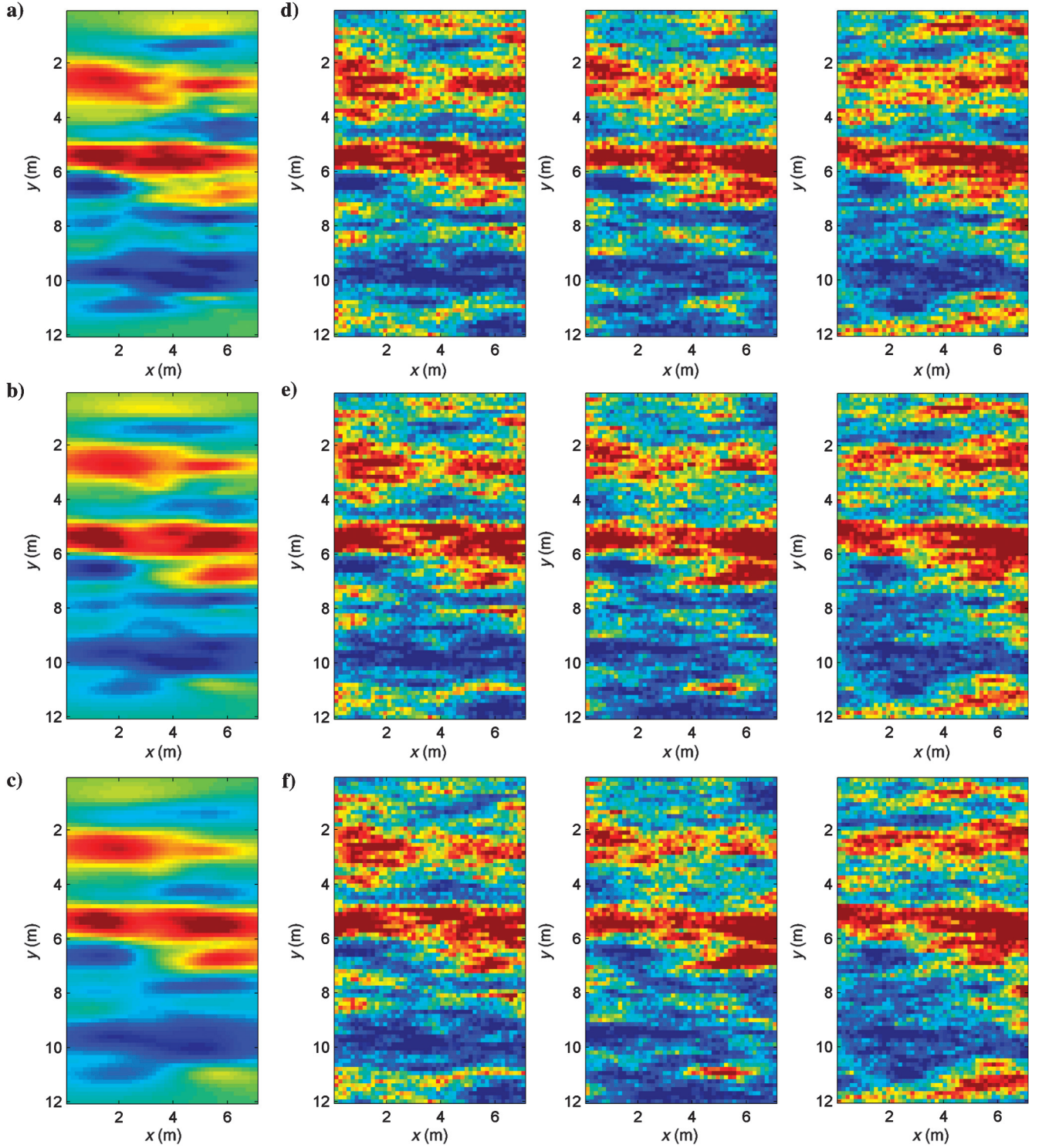


Figure 11. Inversion accounting for estimated modeling error  $C_{T_{app}}$ . (a-c) Maximum a posteriori model (least-squares mean estimate) for  $C_d = 10, 10.1^2$ , and  $10.4 \text{ ns}^2$ . (d-f) Three realizations from the a posteriori probability density considering  $C_d = 10, 10.1^2$ , and  $10.4 \text{ ns}^2$ . Compare to Figure 8 to evaluate the effect of considering modeling error.

individual model parameters. It is apparent that the “image” of posterior probability of a channel, provides a sharper image when disregarding the modeling error (see Figure 16a) than when accounting for the modeling error (Figure 16b). Figure 16c and 16d highlights apparently well-resolved areas where the channels exist with a posterior probability higher than 0.999 (blue) and where a channel does not exist with a posterior probability higher than 0.999 (red), in case disregarding (Figure 16c), and accounting for (Figure 16d), the modeling error. Thus, the areas highlighted in Figure 16c–16d represents model parameters that appear to be close to completely resolved by the inversion. It is clear, that the location of channels are apparently better resolved disregarding the modeling error because larger areas appear well resolved in Figure 16c, as compared to the case of accounting for the modeling error (Figure 16d). However, by comparing Figures 15b and 16 to the reference model, Figure 13, it is clear that when disregarding modeling error, one locates areas with high probability of a channel (and high probability of no channel) that are inconsistent with the reference image. For example, at location  $(x, y) = (3 \text{ m}, 4 \text{ m})$  in Figure 16a, the posterior probability of a channel is near zero. Yet, we know from the reference image that a channel is present. On the other hand, when the modeling error is accounted for through the use of the estimated Gaussian

model  $\mathcal{N}(\mathbf{d}_{T_{\text{app}}}, \mathbf{C}_{T_{\text{app}}})$  the apparent inconsistencies between the posterior samples and the reference model are significantly reduced. Most all the features in Figure 16b and 16d that appear to be well resolved, correspond to real features in the reference image.

Note that the Markov chain Monte Carlo calculations spent 200 times more forward calculations (120,000 in all) than the 600 used for the calibration of the modeling error covariance and bias model. One forward computation takes around 67 s using  $\mathbf{g}_{\text{FW}}$  and 0.05 s using  $\mathbf{g}_{\text{BR}}$ . The total actual computation time for the setup considered here, using the  $\mathbf{g}_{\text{BR}}$  forward model, is  $\sim 13 \text{ h}$  ( $\sim 11 \text{ h}$  for sampling of the modeling error, and 2 h for sampling the solution to the inverse problem). The same 120,000 iterations would take around three months to complete if  $\mathbf{g}_{\text{FW}}$  was used as the forward model, indicating a computational speedup of around 166 using  $\mathbf{g}_{\text{BR}}$  along with a modeling error, compared to using  $\mathbf{g}_{\text{FW}}$  directly. Thus, the use of fast approximate forward models, while at the same time quantifying the associated modeling error, has the potential to allow the use of sampling algorithms with dramatically decreased computational demands.

This second case study illustrates that quantifying the modeling error using a Gaussian model enables sampling of the a posteriori probability density for a nonlinear non-Gaussian inverse problem

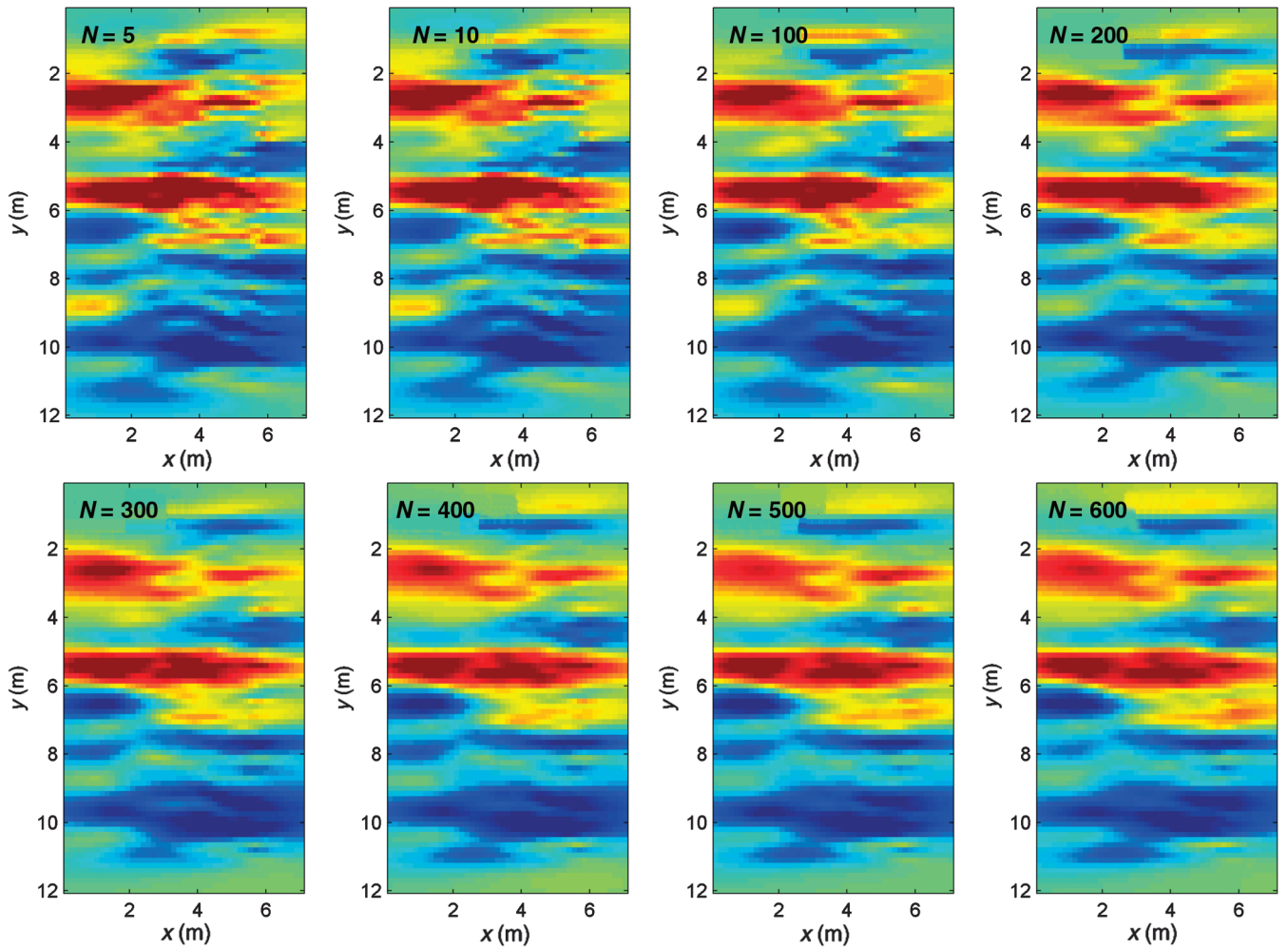


Figure 12. Maximum a posteriori model (least-squares mean estimate) for  $\mathbf{C}_d = \mathbf{I}_0$ , using a sample size of  $N = 5, 10, 100, 200, 300, 400, 500, 600$  to compute  $\mathcal{N}(\mathbf{d}_{T_{\text{app}}}, \mathbf{C}_{T_{\text{app}}})$ .



using an approximate, and fast, forward model in a manner consistent with errors introduced by using the approximate forward model. This has the potential to allow for using fast forward solvers of sampling algorithms such as the Metropolis sampler (Hastings, 1970; Mosegaard and Tarantola, 1995) without introducing unwanted noise mapping errors related to using an approximate forward model.

Cases A and B have demonstrated that disregarding modeling error can introduce features in the a posteriori probability density that appear well resolved, but that are in reality just an effect of mapping the modeling error into the posteriori probability density, as apparently well-resolved features. It will not be possible to quantify which features are consistent with the true subsurface, without solving the inverse problem using the exact forward model with no associated modeling error. By accounting for the modeling error, the resolution of the posterior sample is reduced, compared to disregarding the modeling error, but features that appear resolved, appear to be consistent with the subsurface structures.

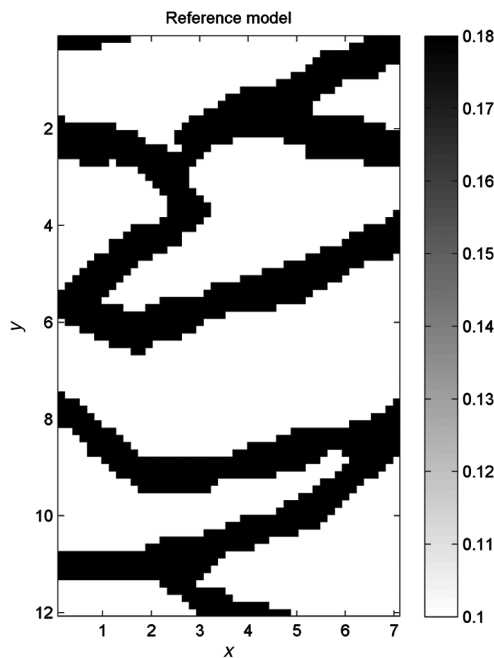


Figure 13. Reference model, case B.

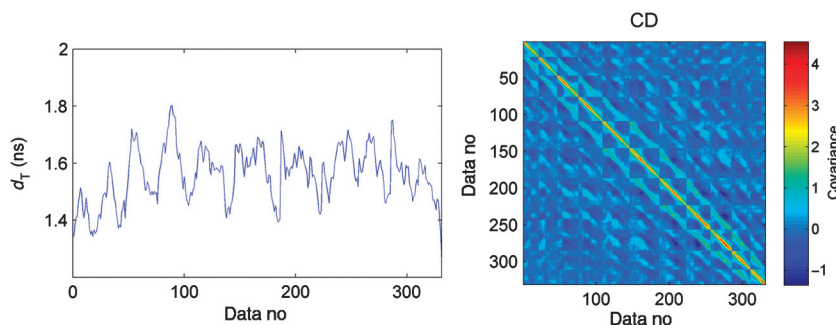


Figure 14. Estimated Gaussian distributed modeling error. (Left)  $d_{T_{app}}$  and (right)  $C_{T_{app}}$

## DISCUSSION

### Computational efficiency

When using sampling methods (e.g., Hastings, 1970; Mosegaard and Tarantola, 1995) to sample the a posteriori distribution, where millions of forward models need to be evaluated, the extra cost of simulating the modeling error may be negligible or at least feasible. For other types of inverse problems, such as, e.g., flow modeling in oil reservoirs, where only a limited number of forward calculations can be evaluated due to computational demands, our approach may be less feasible. Still, if a simple analysis, based on a few models, comparing the forward response of a “perfect” forward model, and one intended to use as the base of an inversion, reveals a significant modeling error (as compared to the measurement uncertainty), our conclusion is clear: Either one should make the effort of trying to quantify the modeling error or, one should not invert the data using the approximate forward model at all because there will be no way of ensuring that modeling noise will not be mapped into the solution of the inverse problem.

An obvious way to reduce the problem of modeling error, if not accounted for during the inversion, is to improve the forward model, if this can be done with little extra computational costs.

The method we have used here to simulate modeling error related to crosshole tomography is, in principle, simple and straightforward. It could, however, have been performed more efficiently if we had used the assumption of stationarity of the a priori model. The traveltimes data obtained with the same relative location to each other will have the same modeling error. For example, the modeling error related to the traveltimes between a source at (1 m, 1 m) and a receiver at (6 m, 5 m) will be identical to the modeling error related to a source at (1 m, 5 m) and a receiver at (6 m, 9 m).

Figures 9 and 14 indicate that the inferred bias and covariance model describing the modeling error possess some smoothness related to the recording geometry used; i.e., data whose sensitivity kernels are located close to each other tend to have a high model error covariance, and those far from each other have a close to zero model error covariance. This suggests that some approximate estimate of the modeling error could be obtained from a model error covariance obtained from a subset of the recorded data. We leave it for future research to explore more efficient ways of characterizing the modeling error.

### Noise versus resolution

We have argued that when a Gaussian distributed modeling error, related to the use of an approximate forward model, can be inferred and is valid, then one can solve the associated inverse problem, without mapping the errors related to modeling error into the solution of the inverse problem. To be very clear: This does not mean that one can obtain an inversion result using an approximate forward model while accounting for modeling errors equivalent to using the full forward model. The noise model used with the approximate forward model will have a higher magnitude (caused by the modeling error) than using the full forward model. When the noise level increases, the resolution will decrease. Thus, one will not

be able to obtain the same resolution using the approximate forward model as when using the full forward model. However, as stressed before, features that appear resolved using the approximate forward model while accounting for the modeling error will be consistent with the true subsurface, if the noise model has been adequately described.

### Source of modeling error

In this study, we propose a general approach to generate a sample of the modeling error. We have applied the methodology to account for imperfect forward models, but the method can be generalized for any modeling error for which the cause of the modeling error can be described by a probability distribution from which realizations can be generated. To use the method, one must be able to describe the source of the modeling error and compute the forward response with and without the source of the modeling error. In the cases investigated here, the source of the modeling error is related to the a priori model, from which realizations are easily generated using geostatistical simulation algorithms.

### Correlated errors in tomographic data

Peterson (2001) and Cordua et al. (2008) list various sources of correlated data errors in tomographic traveltime data. The present study contributes further to the explanation of the correlated data errors often observed in tomographic data by describing a considerable source of correlated data errors, namely, the source that originates from modeling errors related to inadequacies related to the forward problem. This source of errors has not previously been described in the work by Peterson (2001) and Cordua et al. (2008, 2009).

The method presented here can be used to estimate modeling errors stemming from borehole cavities such as those investigated by Cordua et al. (2008, 2009). The requirements are that one must be able to describe borehole cavities in a statistical model and one needs to be able to generate realizations from such a statistical model such that the associated data errors can be simulated (by comparing the forward results of models with and without borehole cavities). In a similar manner, uncertainties about the geometry of the sources and receivers can also be considered and quantified as a

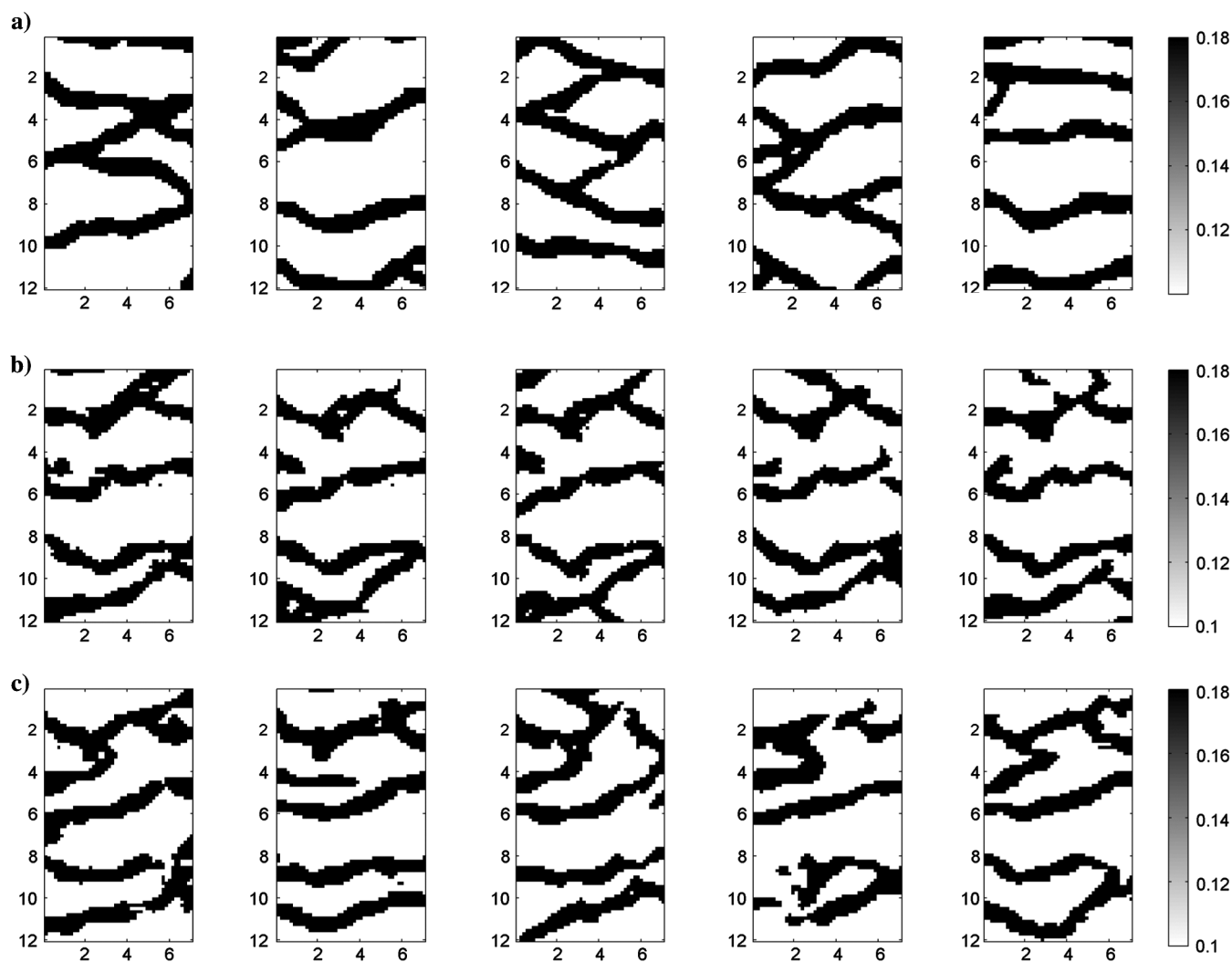


Figure 15. Five independent realization from the (a) prior, (b) posterior disregarding modeling error, and (c) posterior accounting for modeling error.



modeling error, if a statistical model describing such uncertainty can be formulated.

### The Gaussian assumption of the modeling error

We specifically make use of a Gaussian pdf to describe modeling errors. This is convenient because it allows accounting for the modeling error through addition of the covariance models describing measurement uncertainty and modeling error. If another type of parametric probability distribution, as, for example, a Poisson distribution, better describes the simulated modeling errors  $\mathbf{D}_\Lambda$ , such a model should be used instead. The evaluation of the likelihood, equation 6, then becomes more complex.

### Applications

The main application of the method we have proposed is that it allows the solution of an inverse problem, using an imperfect

forward model, without introducing unwanted artifacts in the solution to the inverse problem, as demonstrated in cases A and B.

If the prior model and the noise model are Gaussian, then one can in principle solve such a nonlinear Gaussian inversion problem, using simple linear least-squares inversion. One needs to use an approximate linear forward model and check whether the computed sample of the modeling error can be described by a Gaussian distributed modeling error. The posterior resolution of the model parameters will decrease when using the approximation as opposed to when solving the full nonlinear inverse problem, with no modeling error. This was shown in case A.

Monte Carlo sampling methods can be computationally very expensive (see, e.g., Cordua et al., 2012). Therefore, one may want to make use of approximate forward models. Without accounting for the associated modeling errors, artifacts may be introduced into computed posterior realizations, and the sampling problem may become hard. The method presented here allows accounting for, at least to some degree, the modeling error caused by using an approximate forward model, as demonstrated in case B. This has the potential to allow much faster, but less accurate, forward models to be used as part of more efficient Monte Carlo-based sampling methods.

Another possible application, not demonstrated here, is to make use of an approximate fast forward model to compute an approximation of the true likelihood function. It has been known for many years that sampling of a probability distribution can be made more efficient if we have an easy-to-sample approximation to the distribution we wish to sample. This was exemplified by Mosegaard and Hansen (2007) who demonstrate the computational efficiency obtained using approximations with a rejection sampler. By quantifying the modeling error as proposed in this paper, any simple forward model may in this way act as an approximation to the full forward problem and allow for a potentially much faster sampling algorithm. It is important to note that when using such an approach, where the full forward model is only evaluated when allowed by the approximate and fast likelihood estimate, no reduction in resolution of the a posteriori probability density will appear. The a posteriori probability density (related to the hard inverse problem) will be sampled, as well as if using only the true forward model, but due to the use of approximations, the computational efficiency will be increased.

### CONCLUSIONS

We have proposed a general way to simulate a sample of the modeling error due to the use of an imperfect forward model and/or model description. It can be applied in cases in which a statistical model that describes the cause of the modeling error can be quantified. If a Gaussian model can describe the variability of a sample from the modeling error, such a Gaussian model can be straightforwardly inferred from the sample of the modeling error. This allows accounting for the modeling error by simple addition of the covariance model describing data uncertainty and modeling error.

We have applied the method to inversion of GPR crosshole traveltime data, using approximate forward models. First, we demonstrated that the modeling error due to the use of approximate forward models can be more than an order of magnitude larger than the measurement error. We demonstrated the use of approximate forward models to solve a nonlinear inverse problem, with a Gaussian a priori model, using classical least-squares inversion. We also demonstrated how a nonlinear inverse problem, with a

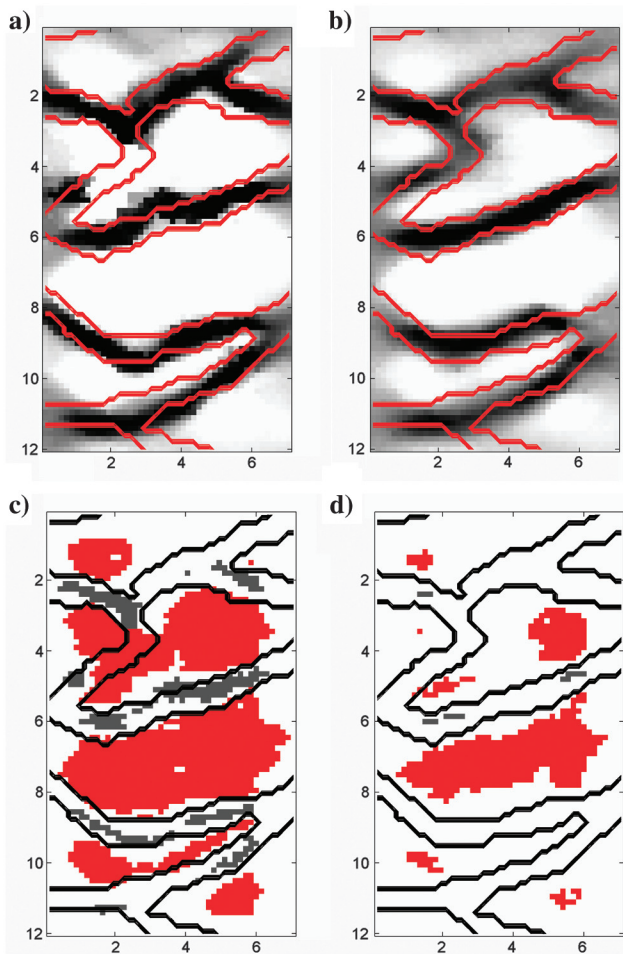


Figure 16. (a-b) A posteriori probability of a channel in case (a) disregarding modeling error and (b) accounting for modeling error. The black color indicates a probability of one, and the white color indicates a probability of zero. The red lines reflect the actual location of channel borders. (c-d) High probability of channel (blue) and high probability of "no channel" in case (a) disregarding modeling error and (b) accounting for modeling error. The black lines reflect the actual location of channel borders.

non-Gaussian a priori model, could be efficiently solved using the extended Metropolis sampler with an approximate, and fast, forward model. In both cases, we found that disregarding the modeling error led to unwanted artifacts in the generated realizations from the a posteriori probability. If not accounted for, the modeling error will be mapped into the a posteriori probability density as unwanted features, that may appear well resolved, but are an effect of fitting noise. On the other hand, when accounting for the modeling error, no apparent unwanted features were noticed in the realizations of the a posteriori probability density. In fact, the features that appeared resolved when accounting for the modeling error were consistent with the features of the known reference model. Finally, we demonstrate that increasing the uncorrelated measurement uncertainty in an attempt to account for the modeling error leads to decreased resolution, while at the same time artifacts appear in realizations of the posteriori probability density.

## ACKNOWLEDGMENTS

We thank M. Bosch, D. Oldenburg, and P. Lelièvre for their constructive, detailed, and insightful comments that helped improve the clarity of the manuscript considerably. This work has been partly sponsored by the Danish Research Council for Strategic Research (journal no. 10-093934).

## REFERENCES

- Bellefleur, G., and M. Chouteau, 2001, Massive sulphide delineation using borehole radar: Tests at the McConnell nickel deposit, Sudbury, Ontario: *Journal of Applied Geophysics*, **47**, 45–61, doi: [10.1016/S0926-9851\(01\)00046-5](https://doi.org/10.1016/S0926-9851(01)00046-5).
- Buland, A., and H. Omre, 2003, Bayesian linearized AVO inversion: *Geophysics*, **68**, 185–198, doi: [10.1190/1.1543206](https://doi.org/10.1190/1.1543206).
- Buursink, M., T. Johnson, P. Routh, and M. Knoll, 2008, Crosshole radar velocity tomography with finite-frequency Fresnel volume sensitivities: *Geophysical Journal International*, **172**, 1–17, doi: [10.1111/j.1365-246X.2007.03589.x](https://doi.org/10.1111/j.1365-246X.2007.03589.x).
- Cerveny, V., and J. Soares, 1992, Fresnel volume ray tracing: *Geophysics*, **57**, 902–915, doi: [10.1190/1.1443303](https://doi.org/10.1190/1.1443303).
- Christiansen, A., E. Auker, and A. Viezzoli, 2011, Quantification of modeling errors in airborne TEM caused by inaccurate system description: *Geophysics*, **76**, no. 1, F43–F52, doi: [10.1190/1.3511354](https://doi.org/10.1190/1.3511354).
- Cordua, K., T. Hansen, and K. Mosegaard, 2012, Monte Carlo full-waveform inversion of crosshole GPR data using multiple-point geostatistical a priori information: *Geophysics*, **77**, no. 2, H19–H31, doi: [10.1190/geo2011-0170.1](https://doi.org/10.1190/geo2011-0170.1).
- Cordua, K., M. Looms, and L. Nielsen, 2008, Accounting for correlated data errors during inversion of cross-borehole ground penetrating radar data: *Vadose Zone Journal*, **7**, 263–271, doi: [10.2136/vzj2007.0008](https://doi.org/10.2136/vzj2007.0008).
- Cordua, K., L. Nielsen, M. Looms, T. Hansen, and A. Binley, 2009, Quantifying the influence of static-like errors in least-squares-based inversion and sequential simulation of cross-borehole ground penetrating radar data: *Journal of Applied Geophysics*, **68**, 71–84, doi: [10.1016/j.jappgeo.2008.12.002](https://doi.org/10.1016/j.jappgeo.2008.12.002).
- Cuffin, B. N., 1981, Effects of modeling errors on least squares error solutions to the inverse problem of electrocardiology: *Annals of Biomedical Engineering*, **9**, 369–382, doi: [10.1007/BF02364657](https://doi.org/10.1007/BF02364657).
- Dahlen, F., S. Hung, and G. Nolet, 2000, Fréchet kernels for finite-frequency traveltimes — I. Theory: *Geophysical Journal International*, **141**, 157–174, doi: [10.1046/j.1365-246X.2000.00070.x](https://doi.org/10.1046/j.1365-246X.2000.00070.x).
- Ernst, J. R., H. Maurer, A. G. Green, and K. Holliger, 2007, Full-waveform inversion of crosshole radar data based on 2-D finite-difference time-domain solutions of Maxwell's equations: *IEEE Transactions on Geoscience and Remote Sensing*, **45**, 2807–2828, doi: [10.1109/TGRS.2007.901048](https://doi.org/10.1109/TGRS.2007.901048).
- Fullagar, P., D. Livelybrooks, P. Zhang, A. Calvert, and Y. Wu, 2000, Radio tomography and borehole radar delineation of the McConnell nickel sulfide deposit, Sudbury, Ontario, Canada: *Geophysics*, **65**, 1920–1930, doi: [10.1190/1.1444876](https://doi.org/10.1190/1.1444876).
- Hansen, T., K. Cordua, M. Looms, and K. Mosegaard, 2013a, SIPPI: A MATLAB toolbox for sampling the solution to inverse problems with complex prior information: Part 1, Methodology: *Computers & Geosciences*, **52**, 470–480, doi: [10.1016/j.cageo.2012.09.004](https://doi.org/10.1016/j.cageo.2012.09.004).
- Hansen, T., K. Cordua, M. Looms, and K. Mosegaard, 2013b, SIPPI: A MATLAB toolbox for sampling the solution to inverse problems with complex prior information: Part 2, Application to cross hole GPR tomography: *Computers & Geosciences*, **52**, 481–492, doi: [10.1016/j.cageo.2012.10.001](https://doi.org/10.1016/j.cageo.2012.10.001).
- Hansen, T. M., K. C. Cordua, and K. Mosegaard, 2012, Inverse problems with non-trivial priors — Efficient solution through sequential Gibbs sampling: *Computational Geosciences*, **16**, 593–611, doi: [10.1007/s10596-011-9271-1](https://doi.org/10.1007/s10596-011-9271-1).
- Hastings, W., 1970, Monte Carlo sampling methods using Markov chains and their applications: *Biometrika*, **57**, 97–109, doi: [10.1093/biomet/57.1.97](https://doi.org/10.1093/biomet/57.1.97).
- Hubbard, S., J. Peterson, Jr., E. Majer, P. Zawislanski, K. Williams, J. Roberts, and F. Wobber, 1997, Estimation of permeable pathways and water content using tomographic radar data: *The Leading Edge*, **16**, 1623–1630, doi: [10.1190/1.1437539](https://doi.org/10.1190/1.1437539).
- Husen, S., and E. Kissling, 2001, Local earthquake tomography between rays and waves: Fat ray tomography: *Physics of the Earth and Planetary Interiors*, **123**, 127–147, doi: [10.1016/S0031-9201\(00\)00206-5](https://doi.org/10.1016/S0031-9201(00)00206-5).
- Jensen, J., B. Jacobsen, and J. Christensen-Dalsgaard, 2000, Sensitivity kernels for time-distance inversion: *Solar Physics*, **192**, 231–239, doi: [10.1023/A:1005238600080](https://doi.org/10.1023/A:1005238600080).
- Lane, J., Jr. F. Haeni, and F. Day-Lewis, 1998, Use of time-lapse attenuation-difference radar tomography methods to monitor saline tracer transport in fractured crystalline bedrock: *Proceedings of 7th International Conference on Ground Penetrating Radar*, University of Kansas, 533–538.
- Lekić, V., and B. Romanowicz, 2011, Inferring upper-mantle structure by full waveform tomography with the spectral element method: *Geophysical Journal International*, **185**, 799–831, doi: [10.1111/j.1365-246X.2011.04969.x](https://doi.org/10.1111/j.1365-246X.2011.04969.x).
- Lelièvre, P. G., C. G. Farquharson, and C. A. Hurich, 2011, Inversion of first-arrival seismic traveltimes without rays, implemented on unstructured grids: *Geophysical Journal International*, **185**, 749–763, doi: [10.1111/j.1365-246X.2011.04964.x](https://doi.org/10.1111/j.1365-246X.2011.04964.x).
- Liu, Y., L. Dong, Y. Wang, J. Zhu, and Z. Ma, 2009, Sensitivity kernels for seismic Fresnel volume tomography: *Geophysics*, **74**, no. 5, U35–U46, doi: [10.1190/1.3169600](https://doi.org/10.1190/1.3169600).
- Looms, M., A. Binley, K. Jensen, L. Nielsen, and T. Hansen, 2008a, Identifying unsaturated hydraulic parameters using an integrated data fusion approach on cross-borehole geophysical data: *Vadose Zone Journal*, **7**, 238–248, doi: [10.2136/vzj2007.0087](https://doi.org/10.2136/vzj2007.0087).
- Looms, M., K. Jensen, A. Binley, and L. Nielsen, 2008b, Monitoring unsaturated flow and transport using cross-borehole geophysical methods: *Vadose Zone Journal*, **7**, 227–237, doi: [10.2136/vzj2006.0129](https://doi.org/10.2136/vzj2006.0129).
- Marquering, H., F. Dahlen, and G. Nolet, 1999, Three-dimensional sensitivity kernels for finite-frequency traveltimes: The banana-doughnut paradox: *Geophysical Journal International*, **137**, 805–815, doi: [10.1046/j.1365-246X.1999.00837.x](https://doi.org/10.1046/j.1365-246X.1999.00837.x).
- McGillivray, P., and D. Oldenburg, 1990, Methods for calculating Fréchet derivatives and sensitivities for the non-linear inverse problem: A comparative study: *Geophysical Prospecting*, **38**, 499–524, doi: [10.1111/j.1365-2478.1990.tb01859.x](https://doi.org/10.1111/j.1365-2478.1990.tb01859.x).
- Minsley, B. J., 2011, A trans-dimensional Bayesian Markov chain Monte Carlo algorithm for model assessment using frequency-domain electromagnetic data: *Geophysical Journal International*, **187**, 252–272, doi: [10.1111/j.1365-246X.2011.05165.x](https://doi.org/10.1111/j.1365-246X.2011.05165.x).
- Molynieux, J., and D. Schmitt, 1999, First break timing: Arrival onset times by direct correlation: *Geophysics*, **64**, 1492–1501, doi: [10.1190/1.1444653](https://doi.org/10.1190/1.1444653).
- Moran, M., and R. Greenfield, 1993, Radar signature of a 2.5-D tunnel: *Geophysics*, **58**, 1573–1587, doi: [10.1190/1.1443373](https://doi.org/10.1190/1.1443373).
- Mosegaard, K., and T. Hansen, 2007, Is seismic waveform inversion a hard problem, in M. Papadarakakis, D. C. Charmpis, N. D. Lagaros, and Y. Tsompanakis, eds., *ECCOMAS Thematic Conference on Computational Methods in Structural Dynamics and Earthquake Engineering*: SPE, 833–842.
- Mosegaard, K., and M. Sambridge, 2002, Monte Carlo analysis of inverse problems: *Inverse Problems*, **18**, R29–R54, doi: [10.1088/0266-5611/18/3/201](https://doi.org/10.1088/0266-5611/18/3/201).
- Mosegaard, K., and A. Tarantola, 1995, Monte Carlo sampling of solutions to inverse problems: *Journal of Geophysical Research*, **100**, 12431–12447, doi: [10.1029/94JB03097](https://doi.org/10.1029/94JB03097).
- Mosegaard, K., and A. Tarantola, 2002, Probabilistic approach to inverse problems, in W. H. K. Lee, P. Jennings, C. Kisslinger, and H. Kanamori, eds., *International handbook of earthquake and engineering seismology*, vol. 81A: Academic Press, 237–265.
- Olsson, O., L. Falk, O. Forslund, L. Lundmark, and E. Sandberg, 1992, Borehole radar applied to the characterization of hydraulically conductive fracture zones in crystalline rock 1: *Geophysical Prospecting*, **40**, 109–142, doi: [10.1111/j.1365-2478.1992.tb00367.x](https://doi.org/10.1111/j.1365-2478.1992.tb00367.x).

- Peterson, J., Jr, 2001, Pre-inversion corrections and analysis of radar tomographic data: *Journal of Environmental and Engineering Geophysics*, **6**, 1–18, doi: [10.4133/JEEG6.1.1](https://doi.org/10.4133/JEEG6.1.1).
- Saenger, E., and T. Bohlen, 2004, Finite-difference modeling of viscoelastic and anisotropic wave propagation using the rotated staggered grid: *Geophysics*, **69**, 583–591, doi: [10.1190/1.1707078](https://doi.org/10.1190/1.1707078).
- Spetzler, J., and R. Snieder, 2004, The Fresnel volume and transmitted waves: *Geophysics*, **69**, 653–663, doi: [10.1190/1.1759451](https://doi.org/10.1190/1.1759451).
- Strebelle, S., 2000, Sequential simulation drawing structures from training images: Ph.D. thesis, Stanford University.
- Strebelle, S., 2002, Conditional simulation of complex geological structures using multiple-point statistics: *Mathematical Geology*, **34**, 1–21, doi: [10.1023/A:1014009426274](https://doi.org/10.1023/A:1014009426274).
- Tarantola, A., 1986, *Inverse problem theory*: Elsevier.
- Tarantola, A., 1988, Theoretical background for the inversion of seismic waveforms including elasticity and attenuation: *Pure and Applied Geophysics*, **128**, 365–399, doi: [10.1007/BF01772605](https://doi.org/10.1007/BF01772605).
- Tarantola, A., 2005, *Inverse problem theory and methods for model parameter estimation*: SIAM.
- Tarantola, A., and B. Valette, 1982a, Generalized nonlinear inverse problems solved using the least squares criterion: *Reviews of Geophysics and Space Physics*, **20**, 219–232, doi: [10.1029/RG020i002p00219](https://doi.org/10.1029/RG020i002p00219).
- Tarantola, A., and B. Valette, 1982b, Inverse problems = quest for information: *Journal of Geophysics*, **50**, 159–170.
- Tromp, J., C. Tape, and Q. Liu, 2005, Seismic tomography, adjoint methods, time reversal and banana-doughnut kernels: *Geophysical Journal International*, **160**, 195–216, doi: [10.1111/j.1365-246X.2004.02453.x](https://doi.org/10.1111/j.1365-246X.2004.02453.x).
- Troncke, J., K. Holliger, W. Barrash, and M. Knoll, 2004, Multivariate analysis of cross-hole georadar velocity and attenuation tomograms for aquifer zonation: *Water Resources Research*, **40**, doi: [10.1029/2003WR002031](https://doi.org/10.1029/2003WR002031).
- Vidale, J., 1988, Finite-difference calculation of travel times: *Bulletin of the Seismological Society of America*, **78**, 2062–2076.
- Viezzoli, A., A. Christiansen, E. Auken, and K. Sørensen, 2008, Quasi-3D modeling of airborne TEM data by spatially constrained inversion: *Geophysics*, **73**, no. 3, F105–F113, doi: [10.1190/1.2895521](https://doi.org/10.1190/1.2895521).
- Woodward, M. J., 1992, Wave-equation tomography: *Geophysics*, **57**, 15–26, doi: [10.1190/1.1443179](https://doi.org/10.1190/1.1443179).